The Use of Sulfur in Aqueous Carbohydrate Chemistry

A thesis submitted in partial fulfilment of the requirements for the degree of **Doctor of Philosophy** in **Chemistry** at the **University of Canterbury**

Stewart R. Alexander

Supervisor: Professor Antony Fairbanks

University of Canterbury

2017

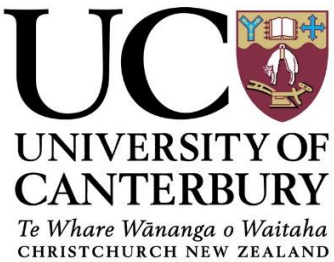


Table of contents

Abstract

The use of sulfur in aqueous carbohydrate chemistry

Carbohydrates play critical roles in a number of biological systems, including cell-cell recognition and signalling, regulation of protein transport, and receptor activation. Carbohydrates are typically found as glycoconjugates: carbohydrates attached to other biomolecules, such as proteins or lipids. Synthetic glycoconjugates have great potential to improve the treatments for a wide range of diseases, and can also be effective vaccines. Carbohydrates and carbohydrate mimics can also interact with a wide range of biological pathways, meaning that techniques for the synthesis of a range of carbohydrate and carbohydrate-derived species are in high demand. However, the modification of carbohydrates and the synthesis of glycoconjugates is often challenging, requiring multiple protection/deprotection steps, and so simple, widely applicable methods for the modification of carbohydrates and the synthesis of glycoconjugates are highly desirable.

Sulfur is a versatile element: it is more nucleophilic than oxygen, can exist in a wide range of oxidation states, and functional groups containing sulfur are capable of carrying out a number of unique reactions including the thiol-ene click reaction, a wide range of redox reactions, and a number of nucleophilic substitution reactions.

The work in this thesis focuses on the application of a number of sulfur's unique features to the preparation of glycoconjugates and glycosidase inhibitors without the use of protecting groups, over a minimal number of steps, in water.

Acknowledgements

Firstly, I would like to thank my supervisor, Professor Antony Fairbanks, for being an amazing lecturer and supervisor throughout my time at UC. You were a great help with everything from research direction, to experimental design, to writing up, and it's been a privilege to be part of your group.

Secondly, I would like to thank the other members of the Fairbanks group. Special thanks goes to Dr. Andrew Watson, for his help when I first started in the lab, and Dr. Davey Lim, for introducing me to DMC, the reagent which later became a core part of my thesis. In addition, thank you to Dr. Akshita Wason for preparing and providing GABARAPL2, and Dr. Vivek Poonthiyil for preparing the sialoglycan. It's also been a pleasure to work with the rest of the Fairbanks group, past and present, including Pragya, Govind, Kajitha, Yusuke, Jesse, Sinthu, Jude, Cormac, and everyone else over the years. You made the lab a great place to work.

I'd like to thank the Parker group for helping me with the biochemical work I carried out, allowing me to use their equipment, and being patient with all my questions.

This research wouldn't have been possible without the technical staff at UC. Dr.s Marie Squire and Amelia Albrett have been invaluable, thank you for all your assistance with NMR, HPLC, and, particularly, mass spectrometry. Thank you also to Dr. Matthew Polson for assistance whenever I came across a new reagent, and for fixing the polarimeter whenever it stopped working. Wayne, Nick, Danny, Gill, Laurie, and Rob, do a great job of keeping the labs running stocked, and again, my research would not have been possible without them.

The Brimble Group in Auckland provided me with peptides for use in my research, for which I am very grateful. Thank you Zaid Anso and Dr. Geoff Williams.

Thank you to UC (Brownlie Scholarship) and the Claude-McCarthy Fellowship for financially supporting me and providing money towards attending conferences during my Ph.D.

Thank you to everyone I met through University Challenge, particularly Ged, Rachel, Liam, Daniel, Catherine, Alexander, Jack, and, of course, Tom. You helped me get through, gave me some amazing experiences, and stopped me from becoming completely lost in my research.

Finally, and perhaps most importantly, I'd like to thank my family for supporting me throughout my Ph.D. I definitely couldn't have done it without you there for me.

Abbreviations

The following abbreviations have been used in this thesis:

General abbreviations

[α]	Specific rotation
ABq	AB quartet (NMR spectroscopy)
Ac	Acetyl
ADMP	2-Azido-1,3-dimethylimidazolinium hexafluorophosphate
Ar	Generic aromatic functional group
ASGPR	Asialoglycoprotein receptor
Bn	Benzyl
Boc	*tert*-Butyloxycarbonyl
BSA	Bovine serum albumin
c	Concentration
°C	Degrees celcius
calcd.	Calculated
CDMBI	2-Chloro-1,3-dimethyl-1*H*-benzimidazolium chloride
CHES	*N*-Cyclohexyl-2-aminoethanesulfonic acid
CuAAC	Copper-catalysed azide-alkyne cycloaddition
d	Doublet (NMR spectroscopy)
δ	Chemical shift (NMR spectroscopy)
Da	Daltons
dd	Doublet of doublets (NMR spectroscopy)
decomp.	Decomposition

dGMIIb	Golgi α-mannosidase II
DMC	2-Chloro-1,3-dimethylimidazolidinium chloride
DMF	*N*,*N*-Dimethylformamide
DMPA	2,2-Dimethoxy-2-phenylacetophenone
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
Endo H	Endo-β-*N*-acetylglucosaminidase from *Streptomyces plicatus*
Endo M	Endo-β-*N*-acetylglucosaminidase from *Mucor hiemalis*
ENGase	Endo-β-*N*-acetylglucosaminidase
EPO	Erythropoietin
eq.	Equivalents
ESI-TOF	Electrospray ionisation-time of flight detector
Et al.	*Et alia*
Et	Ethyl
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
Fmoc	9-Fluorenylmethoxycarbonyl
g	Gram(s)
GABARAPL2	Gamma-aminobutyric acid receptor-associated protein-like 2
Gal	D-Galactose
GalNAc	*N*-Acetyl-D-galactosamine
Glc	D-Glucose
GlcNAc	*N*-Acetyl-D-glucosamine
h	Hour(s)

HIB	*Haemophilius influenza* type B
HILIC	Hydrophilic interaction liquid chromatography
HIV	Human immunodeficiency virus
hυ	Electromagnetic radiation
HPLC	High-performance liquid chromatography
HRMS	High resolution mass spectrometry
Hz	Hertz
J	Coupling constant (NMR spectroscopy)
I	Inhibitor
IC_{50}	Half maximal inhibitory concentration
i.e.	*Id est*
IgG	Immunoglobulin G
IME	2-Imino-2-methoxyethyl
^{i}Pr	Isopropyl
K_a	Acid dissociation constant
K_i	Inhibition constant
K_m	Michaelis constant
K_m^{app}	Apparent Michaelis constant
KLH	Keyhole limpet hemocyanin
L	Litre(s)
LC-MS	Liquid chromatography-mass spectrometry
lit.	Literature
m	metre(s)
m	multiplet (NMR spectroscopy)
m	*meta*

M	Mole(s) per litre
M	Molecular mass ion (mass spectrometry)
Man	D-Mannose
ManNAc	*N*-Acetyl-D-Mannosamine
*m*CPBA	*meta*-Chloroperoxybenzoic acid
Me	Methyl
MHC	Major histocompatibility complex
min	Minute(s)
mol	Mole(s)
m.p.	Melting point
MS	Mass spectrometry
MUC1	Mucin 1
m/z	Mass to charge ratio
υ	Initial reaction rate (enzyme kinetics)
NADPH	Nicotinamide adenine dinucleotide phosphate
υ_{max}	Infrared spectroscopy peak
NMR	Nuclear magnetic resonance
NtMGAM	*N*-Terminal catalytic domain of maltase-glucoamylase
Nu	Nucleophile
o	*ortho*
p	*para*
Ph	Phenyl
Phth	Phthaloyl
pK_a	$-\log_{10}(K_a)$
*p*MP	*para*-Methoxyphenyl

ppm	Parts per million
q	Quartet (NMR spectroscopy)
quant.	Quantitative
R	Generic organic group, unless specified
R_f	Retention factor
RNA	Ribonucleic acid
RT	Room temperature
s	Second(s)
s	Singlet (NMR spectroscopy)
S	Substrate (enzyme kinetics)
sat.	Saturated
S_N2	Bimolecular nucleophilic substitution
t	Triplet (NMR spectroscopy)
^{t}Bu	*tert*-Butyl
tert	*tertiary*
Tf	Trifluoromethanesulfonate
t.l.c.	Thin layer chromatography
TMS	Trimethylsilyl
Tn antigen	α-*N*-Acetyl-D-galactosamine
Tol	Tolyl
Tris	Tris(hydroxymethyl)aminomethane
Trt	Trityl
TTMSS	Tris(trimethylsilyl)silane
UV	Ultraviolet
V_{max}	Maximum initial rate (enzyme kinetics)

VNTR	Variable number tandem repeat

Amino Acids

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cysteine
DHA		Dehydroalanine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
HAG		Homoallylglycine
His	H	Histidine
HPG		Homopropargylglycine
Ile	I	Isoleucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

A note on atom labelling

Carbohydrates have been named in accordance with IUPAC conventions,[1] and numbered according to carbohydrate convention. The monosaccharide components of oligosaccharides are distinguished by a lower case letter, starting at the reducing terminus.

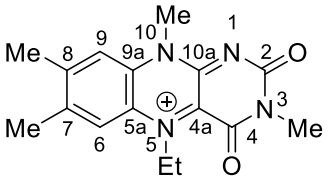


Figure I: Atom labelling for flavins and flavinium salts.

Flavins and flavinium salts have been numbered according to convention,[2] for an example see **Figure I**.

1. McNaught, A. D., *Carbohydr. Res.* **1997,** *297* (1), 1-92.

2. Mishanina, T. V.; Kohen, A., *J. Labelled Compd. Radiopharmaceut.* **2015,** *58* (9), 370-375.

Declaration

I, Stewart Ross Alexander, confirm that the work in this thesis represents my own, that the contribution of any supervisors and others to my research and to the report was consistent with normal supervisory practice, and that any external contributions to the research have been quoted and acknowledged by means of complete references.

Signed: 

Stewart Ross Alexander

Date: 19 July 2017

Chapter 1: Introduction

1.1 Carbohydrates overview

Carbohydrates are a class of biomolecules critical to living organisms.[1] Carbohydrates are found as polymers composed of monosaccharide units, which combine to form oligosaccharides (short carbohydrate chains) and larger polysaccharides (such as cellulose). Unlike other biopolymers, such as DNA, RNA, and proteins, structural diversity in carbohydrates is derived from variation in both the monomers used for construction of the carbohydrate and the position(s) and stereochemistry of chain extension from these monomers. Most monosaccharides have at least three positions from which the carbohydrate chain can be extended, leading to massive potential structural diversity even for short oligosaccharide chains.[2] Modifications such as phosphorylation[3] and sulfation[4] further increase their structural diversity.

1.2 Glycoconjugates

Living organisms take advantage of the structural diversity of carbohydrates through the use of glycoconjugates. Glycoconjugates consist of one or more carbohydrates covalently bonded to another biomolecule, such as a protein or lipid. The structural diversity of carbohydrates leads to a correspondingly diverse set of roles for glycoconjugates.

1.2.1 *N*-Linked glycans

N-Linked glycans are one of the two most common types of protein glycosylation. There are three main types of *N*-linked glycans (**Figure 1.1**), which share a common pentasaccharide core but differ in the sugars attached to this core.[5]

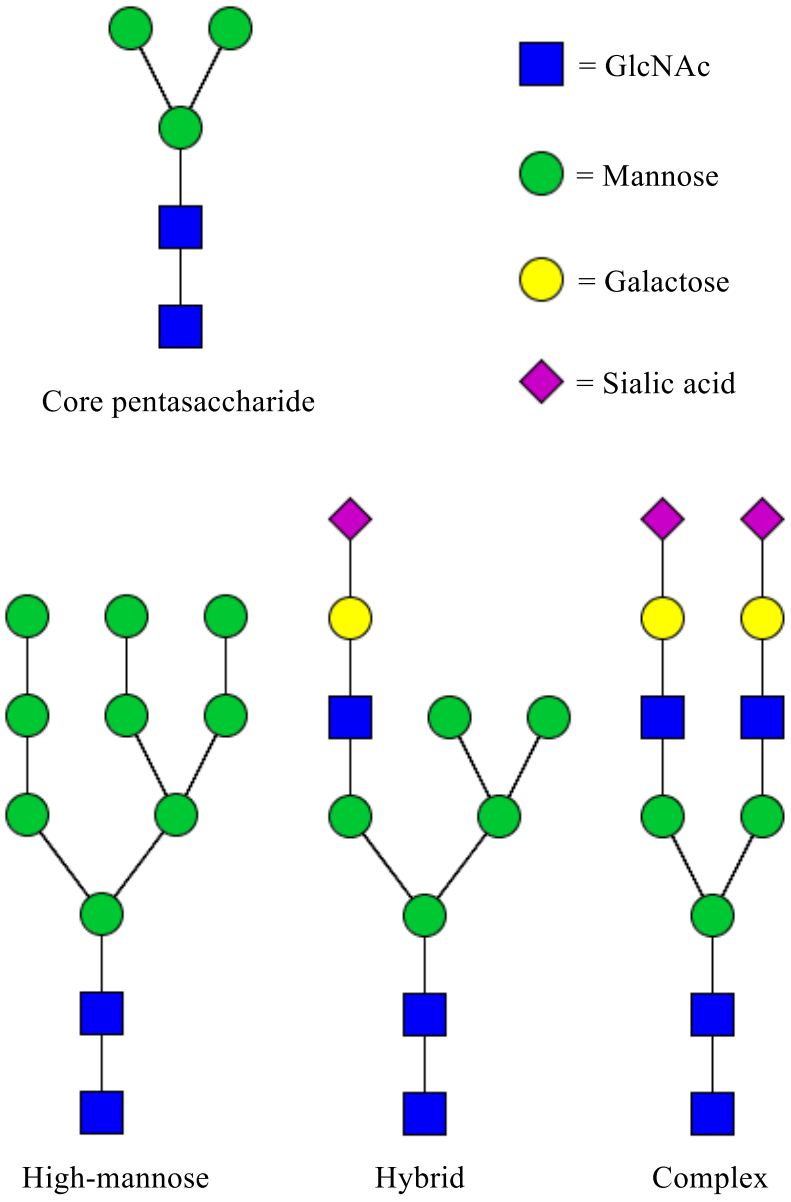


Figure 1.1: Examples of the various types of *N*-linked glycans.

High mannose *N*-linked glycans are composed of the core pentasaccharide with many mannose units attached, while complex *N*-linked glycans have a variety of different sugar units attached to the pentasaccharide core. Meanwhile, hybrid *N*-linked glycans contain one branch of the high-mannose type (which can itself have one or more branches) and one branch of the complex type. *N*-Linked glycans are almost exclusively found attached to protein asparagine residues which match the consensus sequence Asn-X-Ser/Thr, where X is any amino acid except proline.[6] The biosynthesis of *N*-linked glycans begins with the assembly of a $Glc_3Man_9GlcNAc_2$ oligosaccharide attached to dolichol phosphate. The entire oligosaccharide is then transferred to the protein by a glycosyl transferase, and is trimmed and extended as required.[7]

1.2.2 *O*-Linked glycans

Along with *N*-linked glycans, *O*-linked glycans make up the majority of protein glycosylation.[8, 9] *O*-Linked glycans are attached to serine and threonine residues, and unlike *N*-linked glycans only share a common *N*-acetylgalactosamine (GalNAc) residue before diversifying, leading to a greater variety of core structures. *O*-linked glycans are commonly found attached to mucins, proteins characterised by the presence of a repeated section of amino acids, which are typically found as part of mucous secretions and as transmembrane proteins. Unlike *N*-linked glycans, *O*-linked glycans are biosynthesised in a stepwise manner by the sequential addition of monosaccharide units to the growing oligosaccharide. Abnormal *O*-glycosylation has been linked to some forms of cancer.[10]

1.3 Glycoconjugates in protein folding

One of the most important roles of the glycans of *N*-linked glycoproteins is in protein folding.[11] Glycoproteins are typically at least partially dependent on the attachment of appropriate N-linked glycans in order to fold correctly, although the degree to which the glycans are required varies. In extreme situations the lack of glycans on proteins which are normally glycosylated results in misfolding and protein aggregation.[12-14] In other cases the lack of glycans results in an increase in misfolding relative to the glycosylated form[15] or temperature-dependent misfolding.[16] The glycan requirement is not necessarily specific to a single site, rather, one or more of a number of glycans may be required. In this case removal of any individual glycan does not significantly impact protein folding, but removal of multiple glycans can result in misfolding.[13, 17] There is also a subset of glycoproteins in which folding is entirely unaffected by the removal of the associated glycans.[18, 19]

The role of the glycan component of glycoproteins in protein folding is therefore highly variable and unpredictable. The glycans may or may not be essential for correct protein folding, and in the former case all glycans or only some of them may be required.

1.3.1 The calnexin-calreticulin cycle

The calnexin-calreticulin cycle is one of the major pathways by which *N*-linked glycans have an influence on protein folding (**Figure 1.2**).[20, 21]

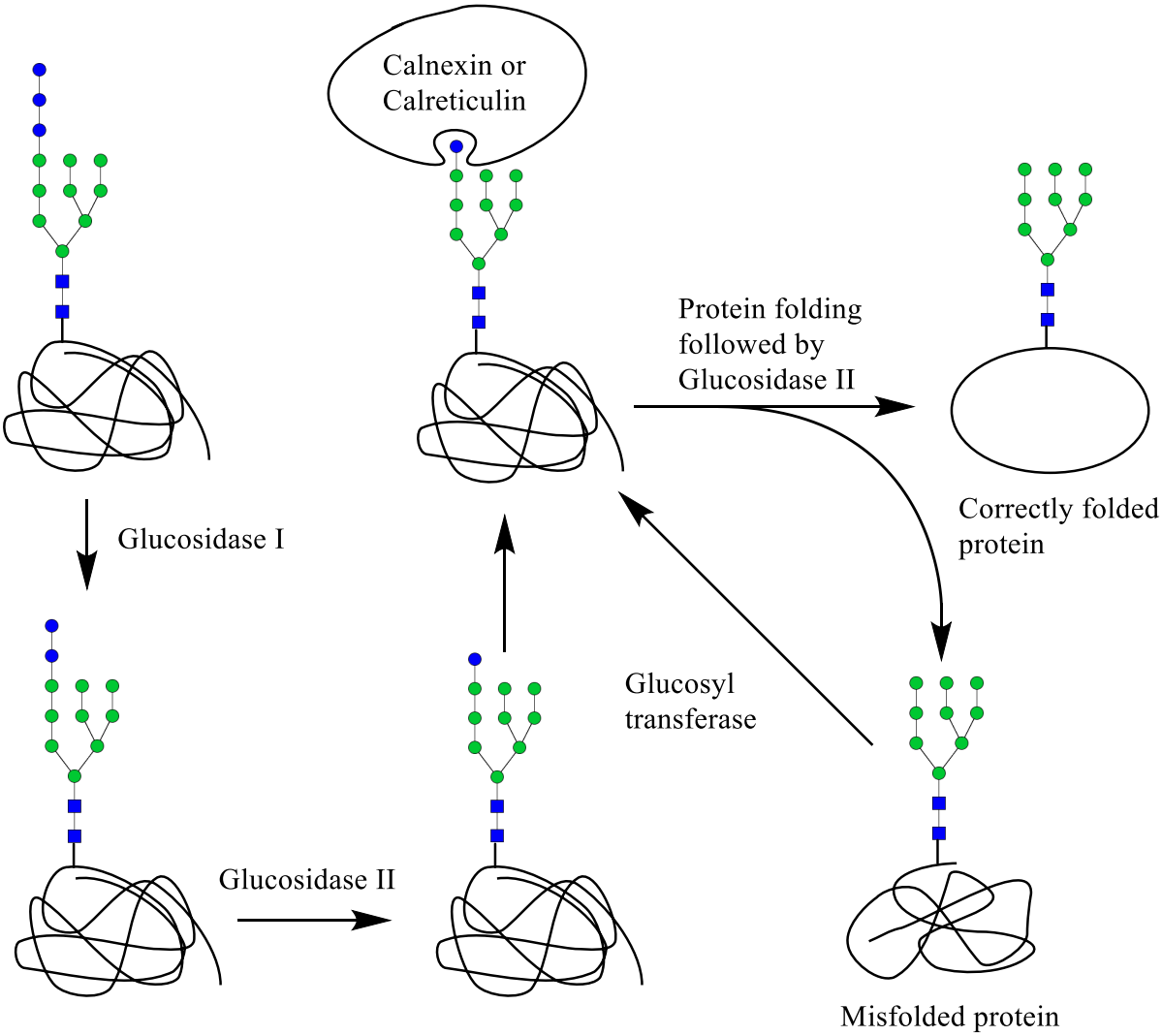


Figure 1.2: The calnexin-calreticulin cycle.

Briefly, calnexin and calreticulin are lectins which bind unfolded or misfolded proteins bearing a monoglucosylated high mannose *N*-linked glycan and act as chaperones to ensure correct folding occurs. Once folding is complete the glucose unit is removed by glucosidase II. However, a glucosyl transferase which recognises misfolded proteins can attach another glucose unit to the free high-mannose *N*-linked glycan, allowing misfolded proteins to re-enter the calnexin-calreticulin cycle in a process which repeats until the protein is correctly folded or degrades.

1.4 Glycan effects on protein stability

As well as assisting in the folding of proteins, glycosylation can impact on the stability of folded glycoproteins, both *in vitro* and *in vivo*. Enzymatic removal of the glycans of glycoproteins or removal of potential glycosylation sites by site-directed mutagenesis can result in a lowering of the melting temperature of the protein, which indicates a decrease in the thermal stability of the protein tertiary structure.[22-24]

For example, the removal of the glycans attached to Asn52 and Asn78 of the α subunit of human chorionic gonadotropin resulted in a small but significant decrease in the melting temperature of the protein.[25]

A similar effect was observed when ribonuclease A was compared to ribonuclease B, which differ only in that ribonuclease B is glycosylated with a high-mannose oligosaccharide at Asn34, while ribonuclease A is not glycosylated.[26] Ribonuclease A and B were found to unfold at similar rates, but ribonuclease B displayed a slightly higher melting temperature.

A study of ovomucoid first domain, a glycoprotein which bears two complex oligosaccharides, revealed that glycosylation increases the melting temperature of the protein, and that this increase is primarily due to entropic effects.[27] The smaller decrease in entropy between the unfolded and folded protein is proposed to be due to an increase in order in the unfolded glycosylated protein, resulting a smaller decrease in disorder when protein folding occurs.

Modification of proteins that are not naturally glycosylated can also result in an increase in the thermal stability of the protein. Glycation of β-lactoglobulin with glucose or fructose *via* the Maillard reaction resulted in an increase in melting temperature of 5 °C and 6 °C respectively.[28] In this case the attachment of carbohydrates to the protein increases the stability of the protein's secondary structure.

In addition to simply increasing the melting temperature of a protein, glycosylation can have a number of other effects on the *in vitro* properties of glycoproteins, including imparting resistance to degradation by proteases,[29] an increase or decrease in protein solubility,[30] and an increase in the protein's resistance to chemically induced unfolding.[31] The glycosylation of proteins is a promising technique for improving the physical properties of proteins.

1.5 Glycan effects on protein transport

The precise glycans attached to a glycoprotein can have an impact on the movement of the glycoprotein *in vivo*. For example, the removal of terminal sialic acid residues from complex *N*-linked glycans results in the rapid clearance of ceruloplasmin by the liver,[32] a process which depends on the binding of terminal galactose residues to the asialoglycoprotein receptor (ASGPR). The ASGPR is responsible for the clearance of a wide range of desialylated glycoproteins,[33, 34] including erythropoietin (EPO). Treatment of EPO with a neuraminidase results in the removal of terminal sialic acid residues from the three complex glycans attached to EPO, and is accompanied by inactivation of EPO *in vivo*,[35] due to the binding of desialylated EPO to the ASGPR and subsequent clearance.[36]

The two mannose-6-phosphate receptors bind selectively to terminal mannose-6-phosphate sugars and are responsible for the transport of a variety of hydrolases and other proteins to the lysosome.[37, 38] Attachment of mannose-6-phosphate sugars to proteins results in enhanced localisation of the enzyme in the lysosome, providing a means of targeting the lysosome in enzyme replacement therapy.[39, 40] The attachment of mannose-6-phosphate is not required to be *via* a natural linkage; localisation is still observed if a non-natural linkage between protein and mannose-6-phosphate is used.[39, 41]

The mannose receptor family[42] is a group of proteins which bind a wide variety of glycans, including sulfated glycans,[43] and glycans possessing terminal mannose or fucose residues.[44] The mannose receptor is responsible for the binding of hormones such as the sulfated glycoprotein hormone lutropin,[45] the clearance of sulfated glycoproteins,[46] and the binding of high-mannose glycoproteins.[47] The mannose receptor also plays a role in the immune system.

1.6 Carbohydrates in self-recognition and immunity

Glycans and glycoproteins are vital components of the immune system.[48] Major histocompatibility complexes (MHCs) are cell surface proteins which bind to glycopeptide antigens and present them for recognition by T cells.[49, 50] There are two classes of MHCs: class I, which present antigens derived from inside the cell, and class II, which present antigens derived from outside the cell. The highest immunogenic activity is observed when both the glycan and peptide antigen components are present.[51] Recognition of a non-self antigen results in an immune response and subsequent generation of memory T cells. Memory T cells cause a rapid immune response upon any further antigen detection, which is reflected in immunity to the pathogen the antigen is taken from.[52]

In addition, the mannose receptor plays a number of roles in the innate immune system.[42] For example, macrophages utilise the mannose receptor to recognise some foreign objects, as human cell surface glycans do not typically possess terminal mannose residues and therefore cells bearing a high number of mannose residues would be expected to be foreign.[53]

1.6.1 Glycoconjugate vaccines

As antigens require both the glycan and peptide component for optimal presentation by MHCs, glycocoproteins or glycopeptides are ideal vaccine candidates.[54, 55]

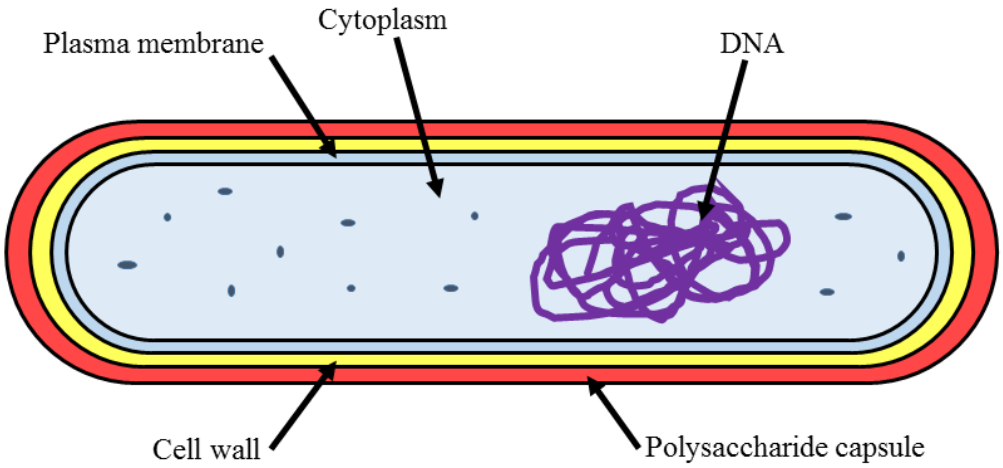


Figure 1.3: Structure of a bacterial cell which produces a polysaccharide capsule

Many bacteria produce a cell capsule constructed from densely packed polysaccharides (**Figure 1.3**), ideal for use as an antigen. The first glycoconjugate vaccine was developed by Robbins and co-workers in 1980 against *Haemophilus influenza* type B (HIB).[56] The vaccine was composed of HIB capsular polysaccharide conjugated to bovine serum albumin (BSA) *via* carbodiimide chemistry, and resulted in a significant immune response, while neither BSA nor polysaccharide alone resulted in any immune response. HIB polysaccharide conjugated to tetanus toxoid is still used as a vaccine against HIB.[57]

A number of other diseases have also been targeted by glycoconjugate vaccines. A cholera vaccine composed of cholera capsule polysaccharide conjugated to cholera toxin displayed significantly higher immunogenic activity than either of the components alone.[58] Similarly, a vaccine against *Shigella flexneri* was prepared by the conjugation of capsule polysaccharide to tetanus toxoid.[59] A vaccine targeting *Streptococcus pneumoniae* composed of 7 different types of capsular polysaccharide fragments, each conjugated to diphtheria CRM_{197} toxin, is effective against pneumonia.[60]

The cell wall of the yeast *Candida albicans* provided another vaccine target.[61] Conjugation of a trimannose component of *Candida albicans'* cell wall to a short peptide resulted in an active anti-fungal vaccine, and conjugation of the same trimannose unit to tetanus toxoid also resulted in an immune response.

As viruses co-opt the biochemical machinery of the host cell in order to reproduce, the glycosylation patterns of viral particles are typically much closer to those of the host, meaning that a carbohydrate-focused vaccine against viruses is much more difficult to produce than for bacteria or fungi.[62] As an example, the human immunodeficiency virus (HIV) is contained within a protein shell which is glycosylated to an extreme degree by high-mannose *N*-linked glycans (**Figure 1.4**).[63] High-mannose *N*-linked glycans are native to the human host, and so do not trigger an immune reaction, and the unique peptide sequences to which they are attached are beneath the thick carbohydrate layer, and so are not detected.

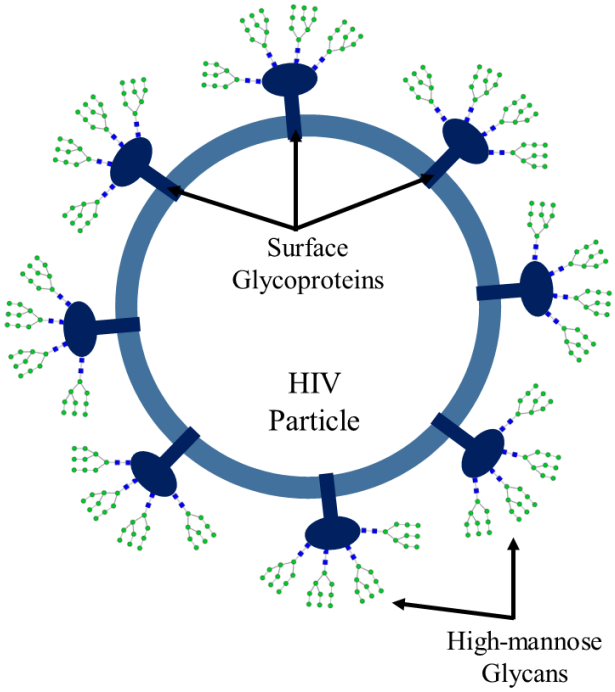


Figure 1.4: External structure of an HIV particle.

Despite these challenges, some progress has been made towards a carbohydrate-based HIV vaccine by mimicking the high-density of high-mannose glycans found on the surface of HIV particles - a feature not naturally found on human glycoproteins. High-mannose clusters based on a galactose support, [64] a short cyclic peptide support,[65] and other branched structures[66-69] showed improved receptor binding and/or immunogenic activity over singly glycosylated structures. The use of viral scaffolds to provide multivalency has also been explored with mixed success.[70, 71] Viruses are clearly more complex vaccine targets than bacteria, however carbohydrate-based vaccines have still shown some potential in provoking an immune response.

In addition to vaccination against pathogens, carbohydrate-based vaccines have also shown some potential as anti-cancer vaccines. Tumour cells often display altered glycosylation patterns relative to non-tumour cells, including increased sialylation, fucosylation, branching

of *N*-linked glycans, and truncation of *O*-linked glycans.[10, 72] Promoting an immune response against these non-self glycans has shown some promise in the development of cancer vaccines. One of the earliest examples was published in 1997 by Ragupathi *et al*.,[73] who demonstrated that the conjugation of globo H, a hexasaccharide associated with some cancers, to keyhole limpet hemocyanin (KLH) resulted in the production of anti-globo H antibodies. Other examples of anti-cancer vaccines based on carbohydrates conjugated to KLH include the sialyl Lewis^y conjugate,[74] the sTn conjugate,[75] and the 5′-*N*-phenylacetyl sTn conjugate,[76] among many others.[77, 78] Many of these vaccines have reached clinical trials.[79]

One feature often observed in tumour cells is the overexpression and aberrant glycosylation of mucins.[80] Mucins are highly *O*-glycosylated proteins found in mucous layers and as transmembrane proteins. Alterations in *O*-linked glycan structure combined with a conserved repeating peptide unit meant that short mucin glycopeptides would be expected to be ideal candidates for a cancer vaccine. Investigation by von Mensdorff-Pouilly and co-workers revealed that this is indeed the case, and that the presence of both the peptide and glycan component resulted in greatly enhanced immunogenic activity.[81] Subsequently a number of investigations into mucin glycopeptide vaccines have been carried out, frequently resulting in promising immune responses.[82-84]

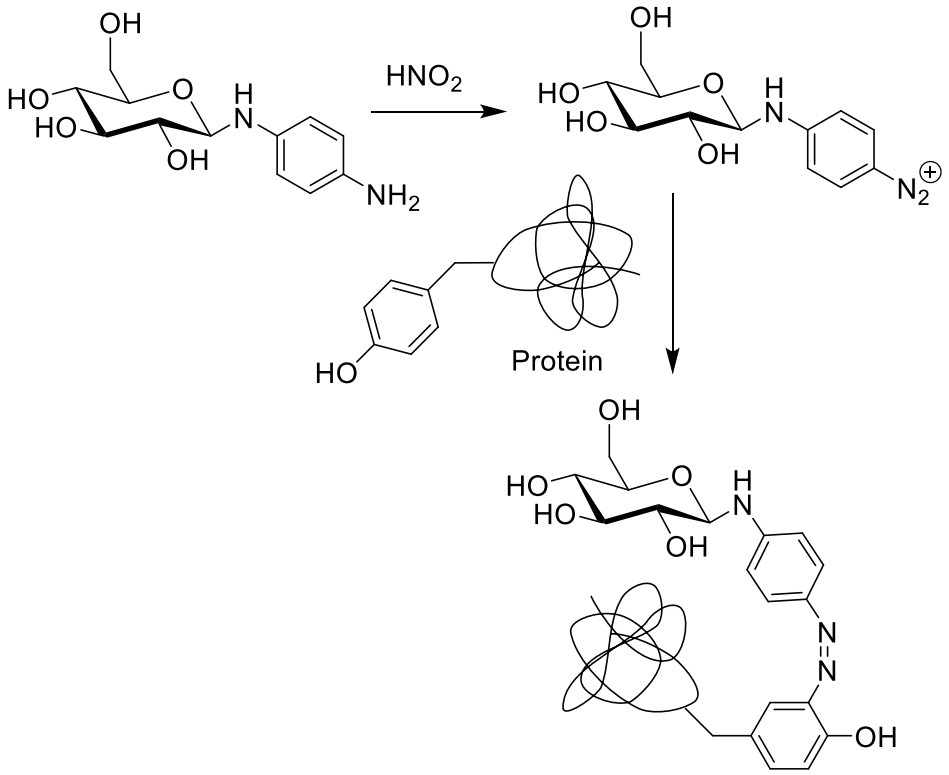
Carbohydrate-based vaccines have shown promise for the immunisation against and treatment of a number of conditions, and further exploration of this field has the potential to yield valuable pharmaceuticals.

1.7 Synthesis of glycoproteins and glycopeptides

The preparation of glycoproteins is a complex field, due to the extremely wide variety of glycans, linkages, and substrates which exist. There are a number of general modifications for the synthesis of glycoproteins, which can broadly be divided into two categories: indiscriminate, where there is no selectivity towards a particular site, and selective, where glycosylation only occurs in one or more pre-selected positions.[85-88]

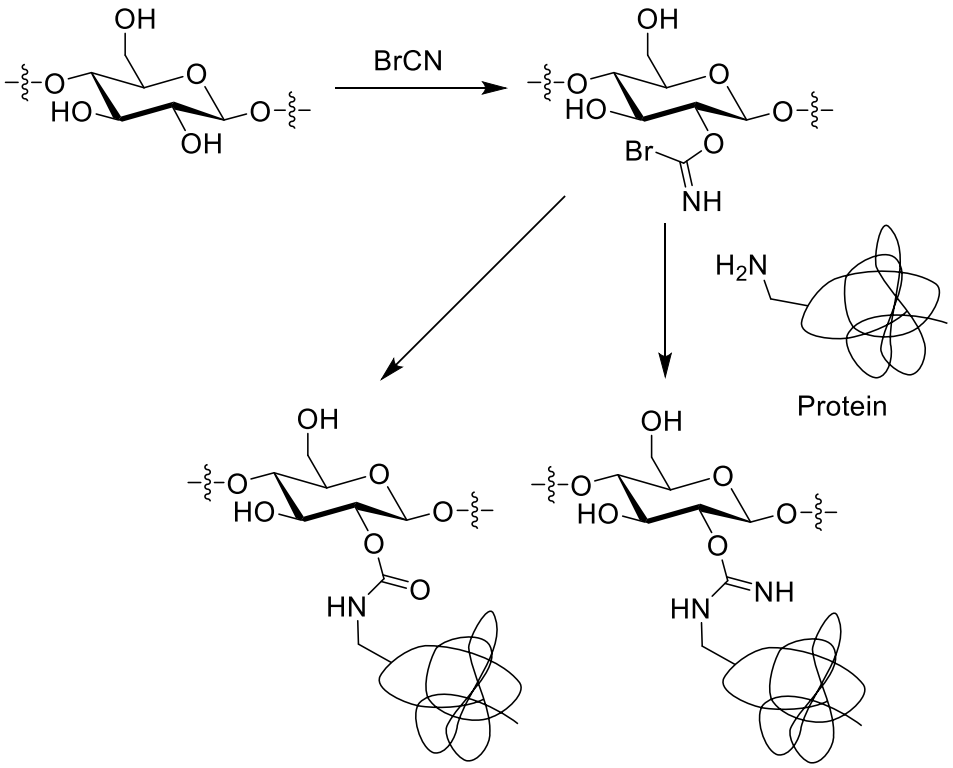
1.7.1 Indiscriminate glycoprotein synthesis

The indiscriminate synthesis of glycoproteins is typically easier to carry out than selective glycosylation, however the glycoproteins produced are also less useful, as glycosylation of each particular molecule is random rather than well defined, which in turn results in more variation in the properties of the glycoprotein between batches. In addition, it is often difficult to control the precise degree of glycosylation, meaning that under- or, more commonly, over-glycosylation, particularly of key catalytic or binding residues, can be an issue. Also, a well-defined structure is necessary for the use of glycoproteins in biomedical applications.[86] Despite these drawbacks, indiscriminate glycosylation methods are more widely applied, due to their relative ease and low cost. The majority of indiscriminate glycosylation methods take advantage of nucleophilic residues, primarily the amine groups of lysine, in order to modify a large number of residues, and are carried out to modify the physical properties (for example, solubility or stability) of the desired conjugate.



Scheme 1.1: Modification of protein with aryldiazonium salts.

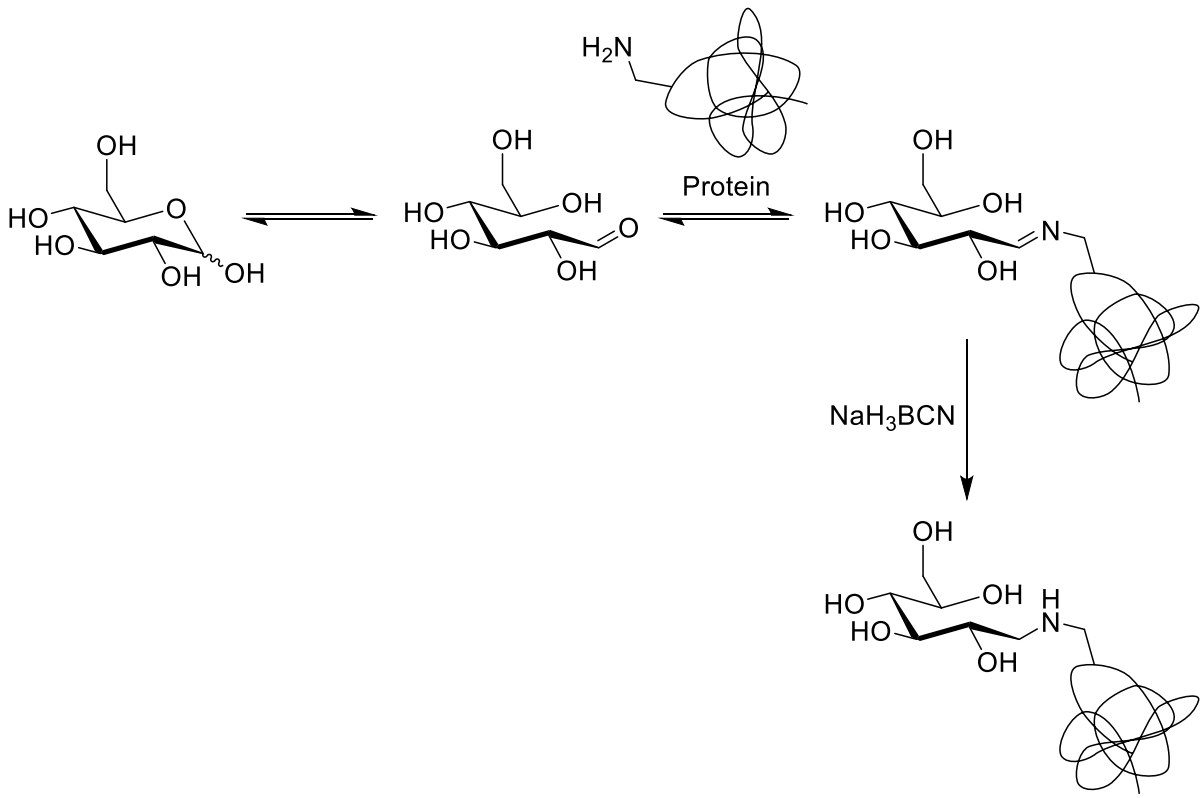
The first reported indiscriminate preparation of a glycoprotein was in 1929 by Goebel and Avery,[89] who prepared the aryldiazonium salts of glucose and galactose and coupled them to horse serum globulin (**Scheme 1.1**). Aryldiazonium salts couple to a wide variety of amino acid side chains, including tyrosine, histidine, and lysine,[90] and therefore provide very little control over the point(s) of attachment of glycans even relative to other indiscriminate glycosylation methods.



Scheme 1.2: The use of cyanogen bromide to prepare glycoproteins.

Cyanogen bromide has been a popular reagent for indiscriminate glycosylation in the past,[91-93] however it has since fallen out of favour due to the harsh conditions required and the hazards associated with the use of cyanogen bromide. In addition, the conjugation reaction is not selective for a single product, resulting in a mixture of possible linkages (**Scheme 1.2**).

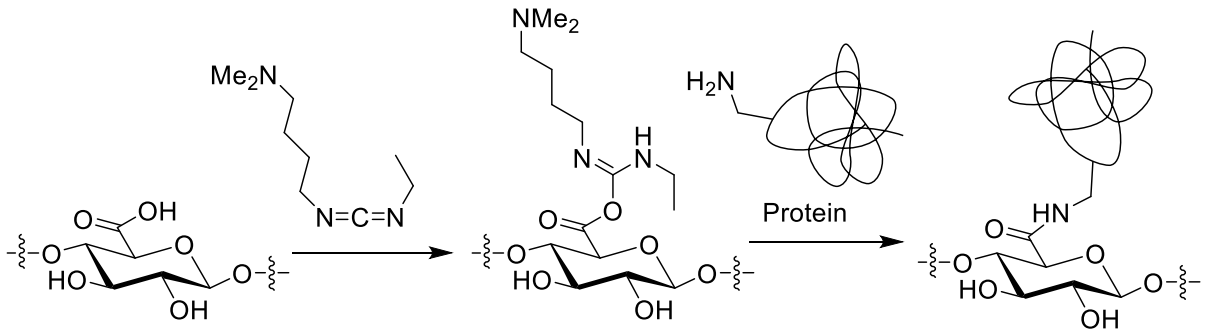
One of the most common procedures for the coupling of carbohydrates to lysine residues is the use of reductive amination. The simplest method of reductive amination relies on the equilibrium which exists between the cyclic hemiacetal and open-chain aldehyde or ketone for all carbohydrates with a free reducing terminus (**Scheme 1.3**).[94-97]



Scheme 1.3: Protein modification by reductive amination.

Under mildly basic conditions the open-chain carbonyl can react in an equilibrium process with the amine group of a lysine residue to form an imine, which is then reduced *in situ* by sodium cyanoborohydride to form a stable secondary amine linkage. Protein modification by reductive amination is ubiquitous, but suffers from the drawback of destroying the natural conformation of the monosaccharide at the reducing terminus of the glycan. Reductive amination may also be carried out at positions other than the reducing terminus of the carbohydrate through the oxidation of other positions, for example by periodate cleavage of sugar rings[98] or by enzymatic oxidation.[99] Glycation followed by Amadori rearrangement shares some features with reductive amination, but notably relies upon the rearrangement of the imine generated as opposed to its reduction. Glycation has found some use in glycoconjugation reactions,[100, 101] however, the glycoconjugates produced by glycation and rearrangement also tend to be

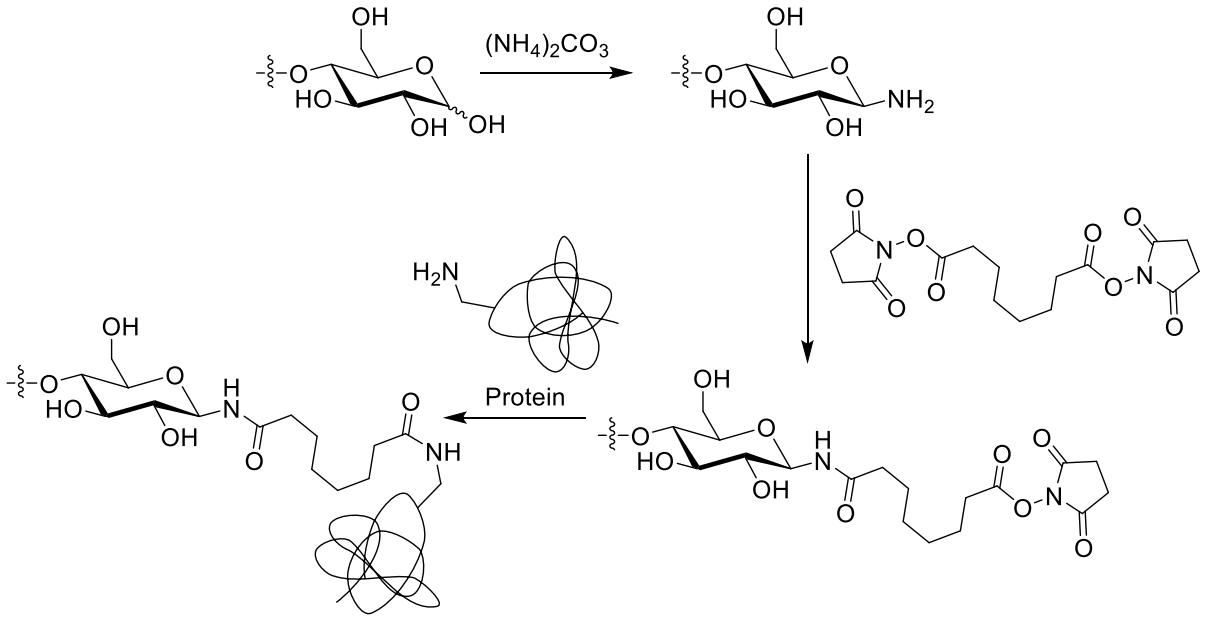
significantly less stable than those produced by reductive amination, and as a result glycation has not seen significant use in the synthesis of glycoproteins.



Scheme 1.4: Carbodiimide-mediated coupling between glyco-carboxylic acids and lysine residues.

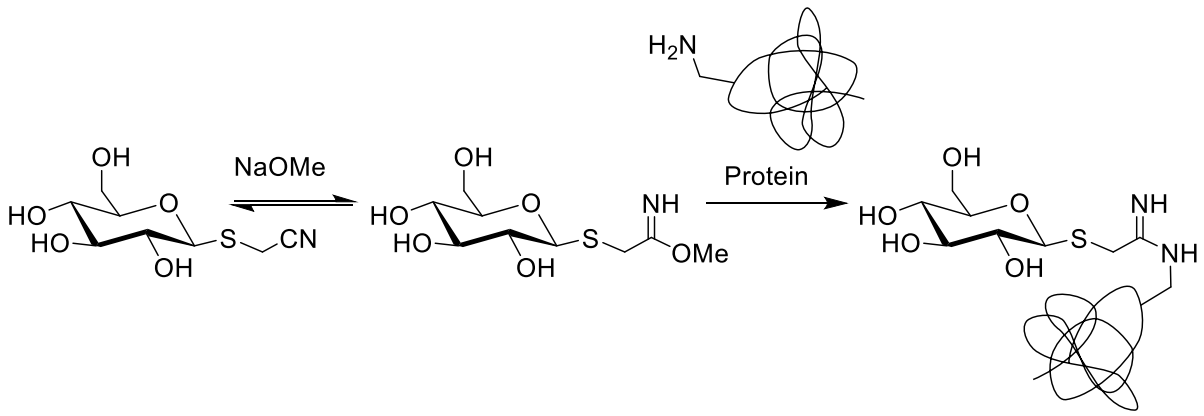
Carbodiimide chemistry has been used to couple carbohydrates to proteins, either *via* the coupling of the carboxylate groups of glucuronic or hyaluronic acid to the lysine residues of a protein (**Scheme 1.4**),[102, 103] or the coupling of amino sugars to the carboxylate groups of aspartic and glutamic acid residues.[104-106] The use of carbodiimide chemistry is limited as either carboxylate or amine functionality is required as part of the sugar, and most natural glycans do not possess these functional groups in appropriate sites for conjugation, limiting the potential carbohydrate substrates.

N-Hydroxysuccinimide esters have also been used to generate glycoproteins through the use of activated esters (**Scheme 1.5**).[107, 108] In this case glycosylamines were accessed through reaction of free reducing sugars with ammonium carbonate, followed by attachment to the lysine residues of a protein through a double-ended succinimide ester linker.



Scheme 1.5: Preparation of glycoproteins using a double-ended *N*-hydroxysuccinimide ester.

2-Imino-2-methoxyethyl thioglycosides (IME thioglycosides) selectively react with lysine residues to produce amidine-linked glycoproteins (**Scheme 1.6**).[109-112] IME thioglycosides are typically prepared by the reaction of cyanomethyl thioglycosides with sodium methoxide in an equilibrium process. The preparation of cyanomethyl thioglycosides is more involved than other indiscriminate glycosylation reagents, requiring the protection of the sugar, conversion to the glycosyl bromide, formation of the glycosyl thiol, reaction of the glycosyl thiol with chloroacetonitrile, and deprotection.[109] The complexity involved in generating the required cyanomethyl thioglycoside can be reduced for synthetic oligosaccharides by introducing the required cyanomethyl functionality during the synthesis, while the glycan target is protected,[113] but IME thioglycosides still represent a more challenging synthetic target than other indiscriminate glycosylating agents, limiting their application. However, IME thioglycosides also represent a more gentle approach to indiscriminate glycosylation than most other methods, maintaining the natural glycan structure.



Scheme 1.6: The use of IME thioglycosides to prepare glycoproteins.

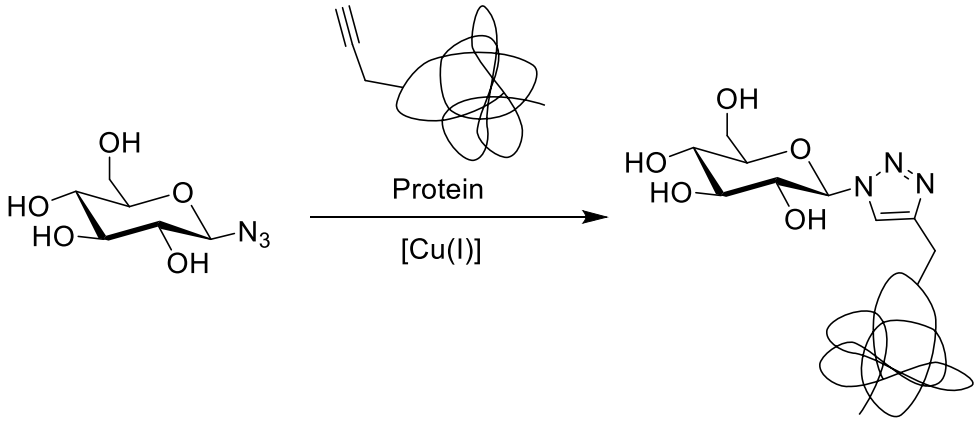
1.7.2 Site-selective glycoprotein synthesis

Well-defined glycoproteins represent more valuable targets than indiscriminately glycosylated glycoproteins. Glycosylation at specific sites is often required for biological activity of glycoproteins, and even in cases where the exact positioning of the glycan is unimportant well-defined glycoproteins are significantly simpler to study, and their synthesis is generally more consistent. However, selectively modifying one or more specified amino acids is often challenging. Site-selective glycoprotein synthesis can be divided into two broad categories: enzymatic glycosylation, which relies on the use of glycosyl hydrolases and transferases to modify existing glycosylation sites,[85] and synthetic glycosylation.[86, 88] Synthetic site-selective glycoprotein synthesis is primarily targeted at amino acids that are not naturally abundant: either proteogenic amino acids naturally found in low concentrations, such as cysteine, or non-proteogenic amino acids.

Enzymatic remodelling of glycoproteins relies upon the vast array of carbohydrate processing enzymes which exist in nature and which can, under the right conditions, be applied to

selectively trim, transfer, and extend the oligosaccharides attached to a protein.[85, 114] A typical procedure begins with either the trimming of the glycan components of the desired protein down to a single monosaccharide using an endohexosaminidase enzyme, such as Endo H,[115, 116] or by total synthesis of the desired protein or peptide, with monosaccharide residues at the desired positions.[117] The desired glycans can be obtained by natural extraction[118, 119] or total synthesis,[40] and then converted to a glycosyl oxazoline. The glycosyl oxazoline and monosaccharide-modified peptide or protein are then used as substrates for a glycosynthase enzyme[120] in order to link the two together. Enzymatic remodelling in this fashion has been used to access a wide range of glycoproteins, including mannosylated glycopeptides for use as vaccines,[121] glycosylated pramlintide for the treatment of diabetes,[117] remodelled IgG antibodies as promising anti-inflammatories,[122] and glycan-modified enzymes for the treatment of lysosomal storage diseases.[123] Enzymatic glycan remodelling has a number of advantages over the chemical synthesis of glycoproteins, including exceptional selectivity, compatibility with a wide range of glycan and protein structures, and often high reaction efficiency. However, enzymatic remodelling is for the most part restricted to natural glycosylation sites and structures, and the activity of glycosynthase enzymes on different proteins and glycans can be variable.

Targeting of non-proteogenic amino acids for the chemical synthesis of glycoproteins offers a number of advantages, most significant of which is that the reactive functional groups, such as alkenes, alkynes, and azides, are not normally found in proteins and peptides, and therefore perfect selectivity can be achieved.



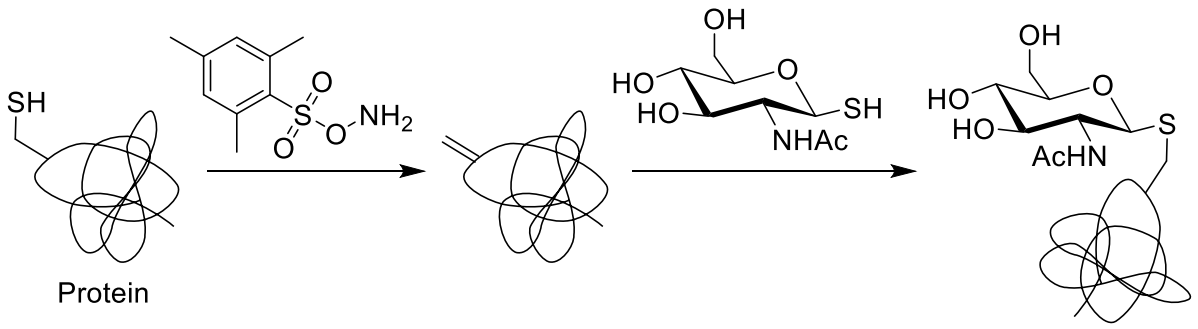
Scheme 1.7: Synthesis of glycoproteins *via* CuAAC.

Copper-catalysed azide-alkyne 1,3-dipolar cycloaddition (CuAAC) is an extremely versatile reaction which couples azides to alkynes selectively, cleanly, efficiently, and under a wide range of conditions.[124, 125] Glycosyl azides have proven to be an effective means of attaching carbohydrates to proteins or peptides through the preparation of peptides or proteins containing unnatural homopropargylglycine (HPG) residues (**Scheme 1.7**).[126-128] CuAAC is an effective means of attaching glycans to proteins with minimal side reactions and high conversions. The primary difficulty in the use of CuAAC to prepare glycoproteins or glycopeptides is the introduction of HPG to the protein or peptide of interest. Peptides can be obtained with relative ease through chemical synthesis,[127] but the incorporation of HPG into proteins is challenging, and relies upon the reassignment of the methionine triplet codon to HPG in protein biosynthesis.[126, 129, 130]

It is also possible to carry out CuAAC to conjugate glycans to proteins in the reverse direction; i.e. with a carbohydrate alkyne and a protein or peptide azide, in a reaction which is functionally very similar to and shares almost all of the benefits and drawbacks of the previously discussed method.[131] Furthermore, the synthesis of a glycan alkyne is in most cases more challenging

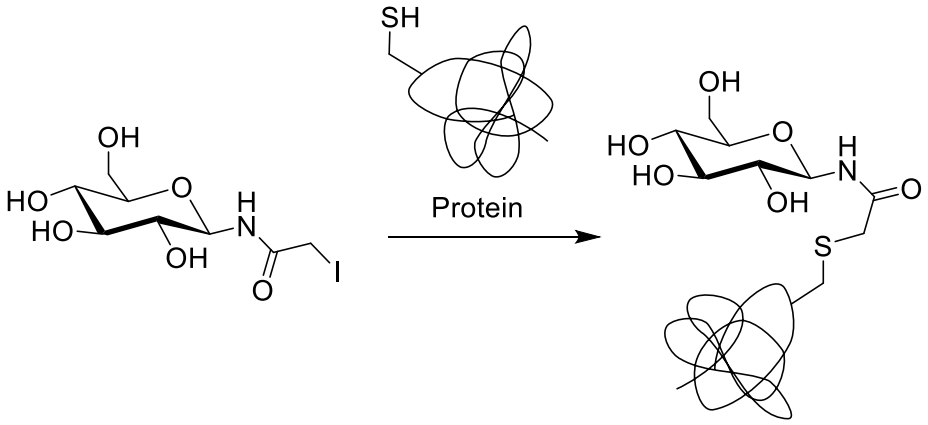
than the synthesis of a glycosyl azide, meaning that the previously discussed method of coupling a glycosyl azide to a protein or peptide alkyne is preferred.

Modification of cysteine residues is one of the leading techniques for glycoprotein synthesis, as cysteine is a proteogenic amino acid and therefore relatively easily introduced into a protein by site-directed mutagenesis, but free cysteine is relatively uncommon in most proteins as it is most often found in disulfide bonds, and therefore undesired side reactions are much simpler to avoid as compared to other proteogenic amino acids.



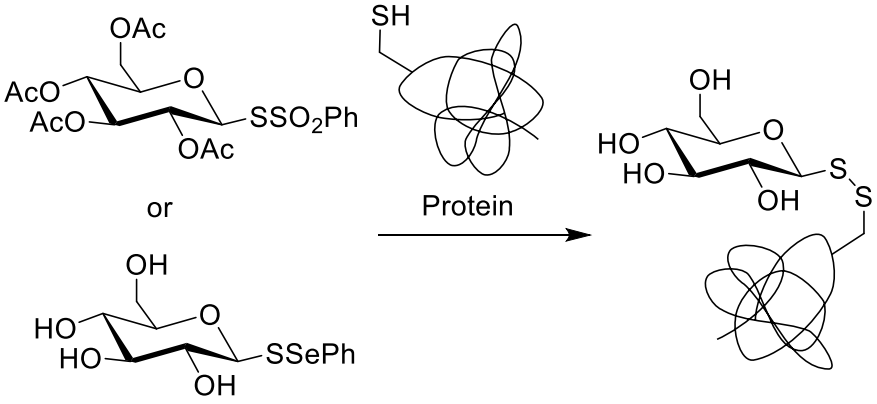
Scheme 1.8: Conversion of cysteine to DHA and reaction with glycosyl thiol to form glycoproteins.

Conversion of cysteine residues to dehydroalanine (DHA) allows for a conjugate addition reaction to take place with an appropriate nucleophile. Bernardes *et al*. have demonstrated that *O*-mesitylenesulfonylhydroxylamine selectively converts cysteine residues into DHA, which then reacts with thiols, including glycosyl thiols, to produce thioether-linked glycoproteins (**Scheme 1.8**).[132] Glycan-DHA conjugates are stable and have similar structures to natural linkages, making them ideal for the preparation of glycoproteins.[126, 133, 134] There are, however, no reports of the preparation of glycan-DHA conjugates larger than disaccharides.



Scheme 1.9: The use of iodoacetamides to prepare glycoconjugates.

Iodoacetamides are another example of a reagent that exploits the increased nucleophilicity of thiols relative to other biological functional groups. Glycosyl iodoacetamides react with cysteine residues to produce glycoproteins (**Scheme 1.9**).[135, 136] Glycoproteins prepared in this way have amide functionality at the anomeric centre, a characteristic of natural *N*-linked glycans which is not often reproduced in synthetic glycoproteins. However, the reaction of iodoacetamides with cysteine residues is comparatively low yielding, meaning it is difficult to purify the resultant mixture of glycosylated and non-glycosylated protein. Side reactions with lysine residues are also possible.[137]

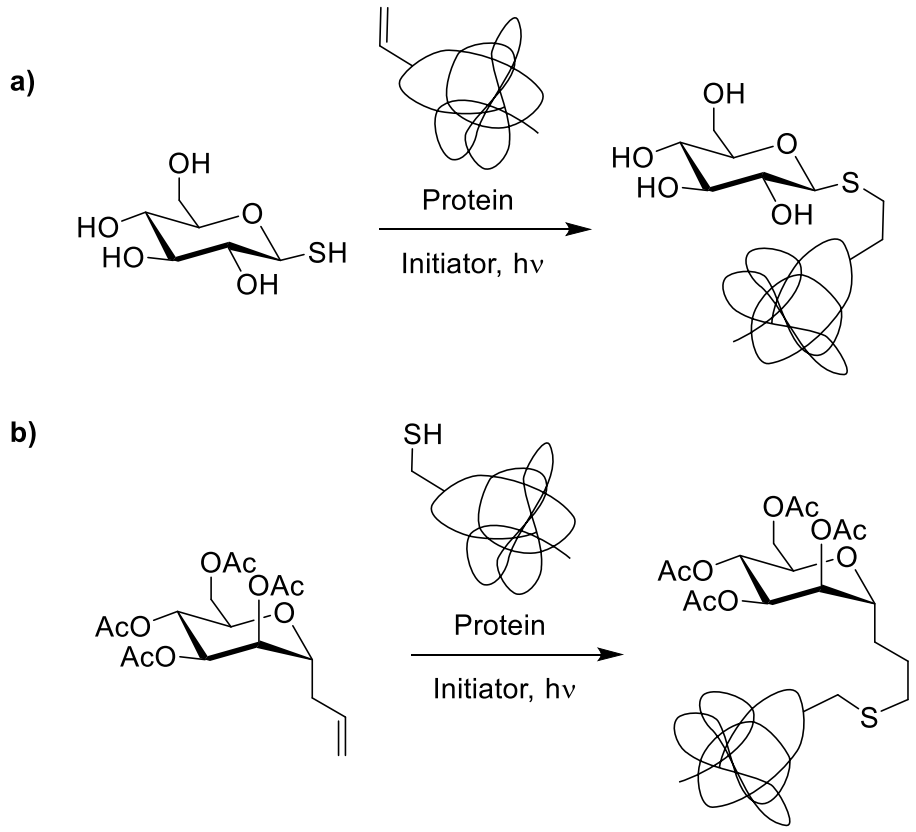


Scheme 1.10: Use of thiosulfonates and selenyl sulfides to prepare glycoproteins.

Cysteine can also be selectively modified through the formation of asymmetric disulfides. A number of reagents have been developed to carry out this transformation, most notably glycosyl

phenyl-[138] and methylthiosulfonates,[139] and glycosyl selenyl sulfides[140] (**Scheme 1.10**). Both rely upon the attachment of a good leaving group to the sulfur of a glycosyl thiol, resulting in a species which reacts selectively with other thiols, such as the cysteines of proteins or peptides, to produce asymmetric disulfides. In the case of the selenyl sulfide method, it is also possible to prepare protein or peptide selenyl sulfide,[141] and indeed this is believed to be the intermediate even when glycosyl selenyl sulfides are used as a reagent. For both thiosulfonates and selenyl sulfides the reaction is perfectly selective for cysteine residues, and the main drawback is that preparation of the glycan component can be challenging for larger oligosaccharides. Glycosyl selenyl sulfides are prepared from a glycosyl thiol, which either requires protection and deprotection of the sugar, or the use of Lawesson's reagent, which is incompatible with 2-acetamido sugars,[141] a major component of most natural glycans. Protecting groups are also required for the preparation of glycosyl thiosulfonates; a method does not currently exist for the preparation of unprotected glycosyl thiosulfonates.

Photocatalysed thiol-ene click chemistry has recently developed as another method for the preparation of glycoproteins.[142, 143] In the presence of an appropriate radical initiator and UV light thiols form thiyl radicals, highly reactive and useful synthetic intermediates.[144] In particular, thiyl radicals react with alkenes and alkynes to form stable thioether linkages. As with CuAAC, there are two possible ways to use thiol-ene click chemistry to generate glycoproteins: either by coupling of a glycosyl thiol and protein or peptide alkene (**Scheme 1.11**, **a)**), or coupling of a sugar alkene to protein or peptide thiol (**Scheme 1.11**, **b)**).



Scheme 1.11: Preparation of glycoproteins by photocatalysed thiol-ene click chemistry.

The first method is generally preferred, primarily because, while synthesis of a glycosyl thiol does require protection and deprotection of the sugar, the difficulty in preparing glycosyl thiols of larger carbohydrates is outweighed by the difficulty of attaching an alkene to the same. As a result, proteins or peptides incorporating alkenes must be prepared. This is accomplished in a similar fashion to the preparation of proteins or peptides containing alkynes; that is, through chemical synthesis or the reassignment of the methionine codon to homoallylglycine (HAG). The reaction of carbohydrate alkenes with protein or peptide thiols is attractive from a protein and peptide synthesis perspective, as it would allow coupling to natural cysteine residues, but currently the preparation of carbohydrate alkenes for larger oligosaccharides provides too great a hurdle. Both methods have been used to prepare glycosylated versions of large proteins. Notably, Davis and co-workers prepared virus particles glycosylated at 180 sites through method **a)**,[145] while Dondoni and co-workers prepared BSA glycosylated with mannose

tetraacetate sugars in three positions *via* method **b)**.[146] As a final note, thiol-ene click chemistry can be extended to alkynes, which allows for two glycosylations at each position.

A wide array of procedures currently exist for indiscriminate and site-selective glycoprotein synthesis, however in most cases one or more flaws or difficulties exist that prevent widespread application. In many cases the key issue is preparation of the reactive oligosaccharide or biomolecule without resorting to extended synthetic procedures or complex biological processes. A method which bypasses both these hurdles would allow for the significantly wider application of glycoproteins and glycopeptides.

1.8 DMC and the aqueous modification of carbohydrates

One of the principal issues in the preparation of glycoconjugates is that where modification of the carbohydrate moiety is required this modification is almost invariably accompanied by protection and deprotection steps. Methods that do not utilise protection and deprotection steps are often characterised by low yields, poor selectivities, or are inapplicable to structurally defined oligosaccharides. Reactions which selectively modify oligosaccharides effectively without requiring protecting groups are therefore highly desirable.

Recently, 2-chloro-1,3-dimethylimidazolinium chloride (DMC, **Figure 1.5**) has emerged as a reagent for the selective modification of the anomeric centre of carbohydrates in water without the use of protecting groups.

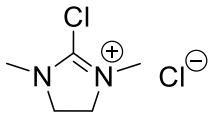
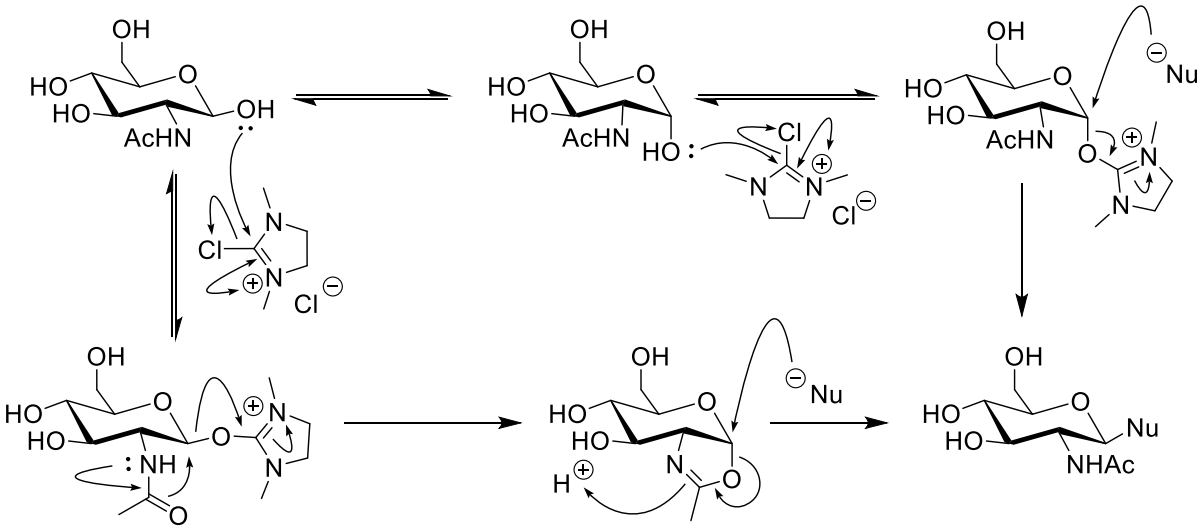


Figure 1.5: 2-Chloro-1,3-dimethylimidazolinium chloride (DMC).

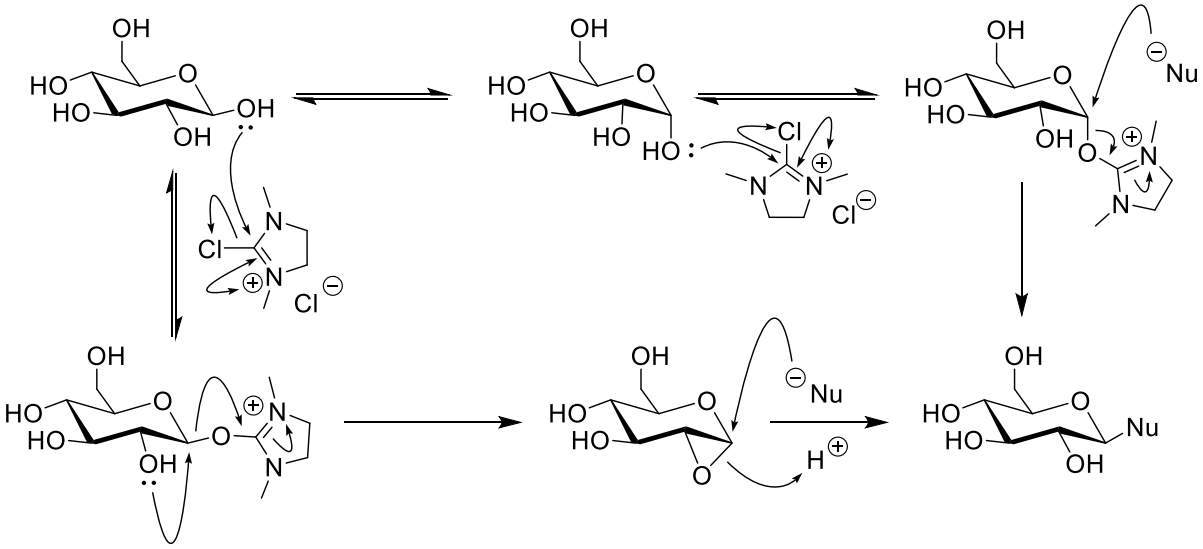
DMC exploits the subtle difference in pK_a between the anomeric hydroxyl group and all other hydroxyl groups of carbohydrates. The anomeric hydroxyl group is more acidic than other oligosaccharide alcohols due to the anomeric effect, and is also more acidic than water.[147] The anomeric hydroxyl group of glucose, for example, has a pK_a of approximately 12, in comparison to values greater than 14 for water and other alcohols. Appropriate selection of base therefore allows the selective activation of the anomeric hydroxyl group in the presence of both other oligosaccharide alcohols and also in the presence of water. The use of a good nucleophile which can out-compete water then allows selective modification of the anomeric centre,[128] to form 1,2-*trans* glycosides. The strength of the nucleophile is critical, as water is present in massive excess and so can potentially outcompete even potent nucleophiles. With this in mind, the use of D_2O as a solvent results in reduced hydrolysis as compared to water, presumably due to the reduced dissociation constant of D_2O.[148]



Scheme 1.12: Activation of 2-acetamido sugars by DMC.

The mechanism of activation by DMC is still not fully understood. Activation of 2-acetamido sugars is the simpler case, and is known to proceed *via* a 1,2-oxazoline, which can subsequently be opened by acid (**Scheme 1.12**).[149] Indeed, DMC has quickly become established as the most widely used reagent for the formation of 1,2-oxazolines from unprotected oligosaccharides comprising 2-acetamido sugars at the reducing terminus.[40, 150]

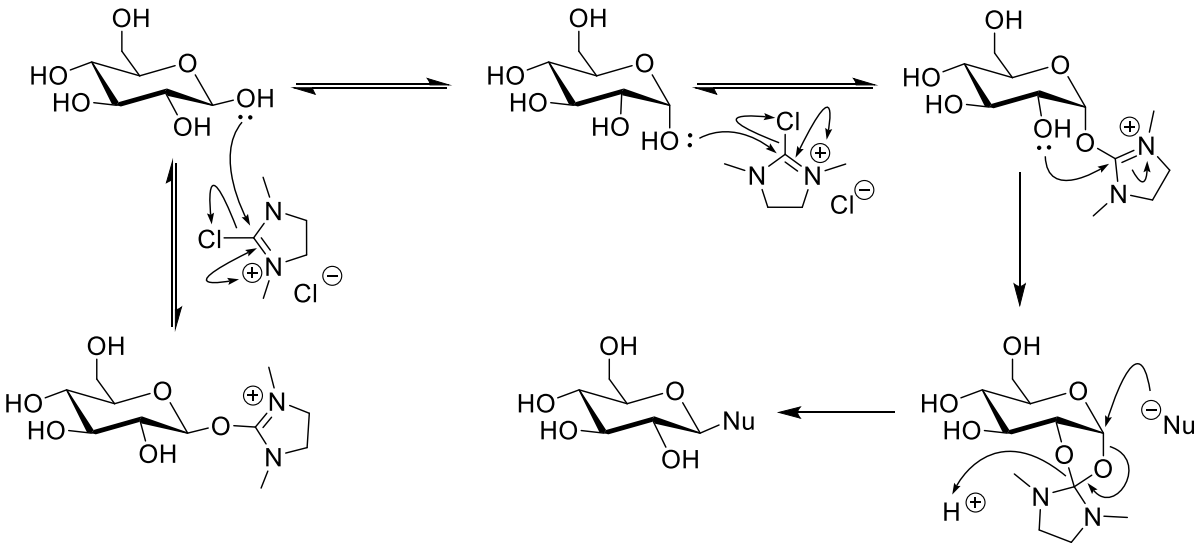
The activation of 2-hydroxy sugars is more complex. In all cases where the sugar contains a 2-hydroxyl group 1,2-*trans* glycosides are formed, with glucose giving primarily β products and mannose giving primarily α products,[151] implying that the 2-hydroxy group is participating in some manner. Shoda *et al*. report that 2-deoxy glucose can be activated by DMC and subsequently allowed to react to produce product, albeit with lowered selectivity, although a yield has not been quoted.[151, 152] Two possible mechanisms have been proposed to account for these observations.



Scheme 1.13: Shoda's proposed DMC activation mechanism.[152]

The first of these mechanisms was proposed by Shoda when he originally developed DMC in 2009 (**Scheme 1.13**).[152, 153] In this proposed mechanism, activation of the β anomer by DMC

results in the formation of a 1,2-anhydrosugar intermediate, which is then opened by a nucleophile to produce the observed product. It is also possible for direct reaction between a nucleophile and the DMC-sugar intermediate to occur, although given that reaction with DMC and subsequent substitution is invariably selective for the 1,2-*trans* product it must be assumed that the rate of reaction for the mechanism involving participation of the 2-hydroxyl group is significantly greater than the rate of direct attack on the DMC-sugar intermediate.



Scheme 1.14: DMC activation mechanism proposed by Novoa *et al*.[154]

An alternative mechanism for activation of sugars by DMC was proposed by Novoa *et al*. (**Scheme 1.14**).[154] Similarly to Shoda's mechanism, the 2-hydroxyl group participates, however, in this case a 5-membered acetal ring is formed, which is then opened by a nucleophile to produce product. In this case the 1,2-*trans* starting anomer cannot possibly react to give product. NMR data purportedly showing the α DMC-glucose intermediate before attack by the 2-hydroxyl group is provided as evidence of this mechanism, however there is minimal characterisation of this intermediate, and the DMC-glucose acetal was not observed.

Upon performing a similar low temperature NMR experiment to Novoa *et al*., Shoda and co-workers unsurprisingly observed the same peaks. However, a more detailed investigation indicated that the peaks observed most likely correspond to 1,2-anhydroglucose,[155] rather than the α DMC-glucose intermediate suggested by Novoa *et al*., implying that DMC activation of a 2-hydroxy sugar results in the formation of a highly reactive 1,2-anhydrosugar, which then reacts selectively with a nucleophile. Although the observation of 1,2-anhydroglucose provides a strong indication of the activation mechanism, a number of mysteries remain, including why alteration of the solvent has a significant effect on the formation of byproducts such as 1,6-anhydroglucose, and why, if direct hydrolysis of the DMC-sugar intermediate is also possible, nucleophilic attack on DMC-activated 2-hydroxy sugars is invariably perfectly selective.

Nevertheless, DMC is undeniably a versatile reagent for the modification of unprotected oligosaccharides. Since the first use of DMC to activate unprotected oligosaccharides a wide variety of nucleophiles have been examined for attack upon the activated intermediate.

As discussed previously, simultaneously one of the most simple and useful applications of DMC is the preparation of 1,2-oxazolines, first demonstrated by Shoda *et al*. in 2009.[149] 1,2-Oxazolines are both useful synthetic intermediates and substrates for endo-β-*N*-acetylglucosaminidase (ENGase) enzymes, which have been widely used in the preparation of glycoconjugates (**Section 1.7.2**). The formation of a 1,2-oxazoline is an inevitable consequence of the DMC activation of 2-acetamido sugars. As the formation of 1,2-oxazoline is currently the most widespread use of DMC activation, attempts to improve this reaction are ongoing.

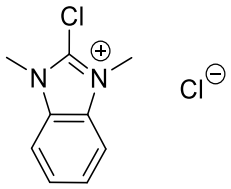
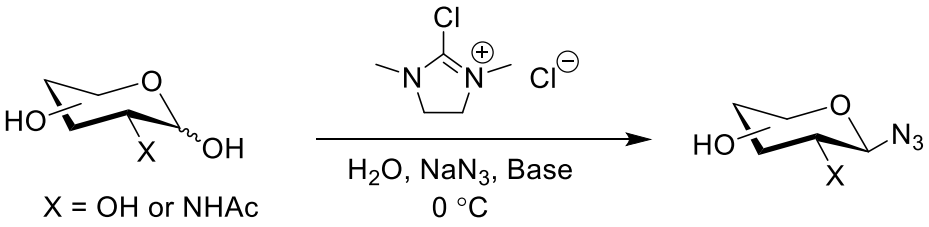


Figure 1.6: 2-Chloro-1,3-dimethyl-1*H*-benzimidazolium chloride (CDMBI).

Notably, 2-chloro-1,3-dimethyl-1*H*-benzimidazolium chloride (CDMBI, **Figure 1.6**) is significantly less hygroscopic than DMC, and the side product of the activation reaction, 1,3-dimethylbenzimidazolone, is almost completely insoluble in water.[156] As a result, the urea byproduct of the reaction can easily be removed by filtration. As the urea product has been reported to inhibit synthetic reactions catalysed by ENGases it therefore becomes possible to use the crude oxazoline as an ENGase substrate, which is not the case with DMC. The most significant issue with CDMBI is that its synthesis is longer than that of DMC, and significantly lower yielding.

Another possible intramolecular reaction that can occur upon DMC activation is the formation of 1,6-anhydrosugar as a result of the attack of the 6-hydroxyl group on the DMC-sugar intermediate.[153] 1,6-Anhydrosugar does not form if 2-acetamido sugars, sugars with an axial 2-hydroxyl group, or sugars without a 6-hydroxyl group are activated by DMC, but in all other cases 1,6-anhydrosugar is invariably observed as a byproduct of DMC-mediated activation.

Strong external nucleophiles can also react with sugars activated by DMC. As discussed previously, the nucleophile must be able to outcompete water or the DMC intermediate will simply hydrolyse, which provides a limit on the nucleophiles that can be used.



Scheme 1.15: Preparation of glycosyl azides by Shoda and co-workers.[151]

A stereotypical strong nucleophile, azide was one of the first nucleophiles to be applied to DMC-mediated substitution reactions (**Scheme 1.15**).[151] Azide proved to be an effective nucleophile, provided a massive excess of sodium azide was used, affording unprotected glycosyl azides in moderate to high yields. The formation of 1,6-anhydrosugar, the main side-product of the reaction, was minimised by changing the base used from triethylamine to diisopropylethylamine or 2,6-lutidine, although in the hands of the Fairbanks group the use of alternative amine bases was never successful. The reaction of 2-acetamido sugars did produce glycosyl azides in high yields, but an extended reaction time (36-48 h) was required, presumably due to the difficulty of opening the 1,2-oxazoline intermediate. Glycosyl azides are useful because they are ideal substrates for CuAAC (**Section 1.7.2**), so a process which gives unprotected glycosyl azides without requiring any other modification or protection of the oligosaccharide substrate is therefore of great benefit for the preparation of glycoconjugates.

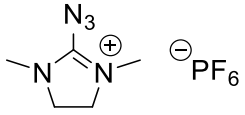
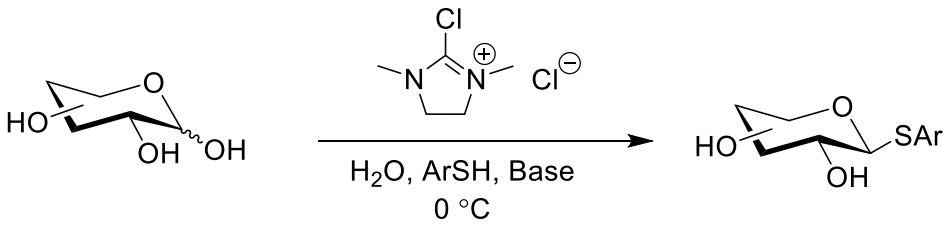


Figure 1.7: 2-Azido-1,3-dimethylimidazolinium chloride, ADMP.

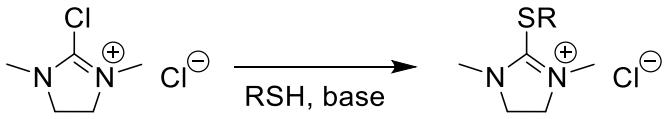
Improving on this method, Lim developed the reagent 2-azido-1,3-dimethylimidazolinium hexafluorophosphate (ADMP, **Figure 1.7**).[128] ADMP generates azide *in situ*, both removing the need to work with highly toxic sodium azide and reducing the number of equivalents of azide required from 10 to 3. In addition, Lim showed that an acidic workup of the ADMP

reaction of 2-acetamido sugars rapidly opens the 1,2-oxazoline intermediate that is formed competitively and forms glycosyl azide product without requiring the extended reaction times used by Shoda *et al*. Furthermore, Lim demonstrated that CuAAC can be carried out in one pot on crude glycosyl azide, providing a rapid means of generating glycoconjugates from a variety of free reducing sugars.



Scheme 1.16: Preparation of unprotected thioglycosides using DMC.

Sulfur-containing species are also widely used for their nucleophilicity, and therefore have also been examined for their reactivity with DMC-activated sugars. Again, the first example was provided by Shoda and co-workers, who demonstrated that DMC activation of 2-hydroxy sugars in the presence of aryl thiols resulted in the selective formation of unprotected aryl thioglycosides (**Scheme 1.16**),[152, 157] species which are useful as *O*-glycoside analogues and, when protected, as donors for glycosylation reactions. In all cases the reaction was very high yielding, although it was not applied to any carbohydrates larger than a disaccharide.



Scheme 1.17: Formation of DMC-thiol adduct.

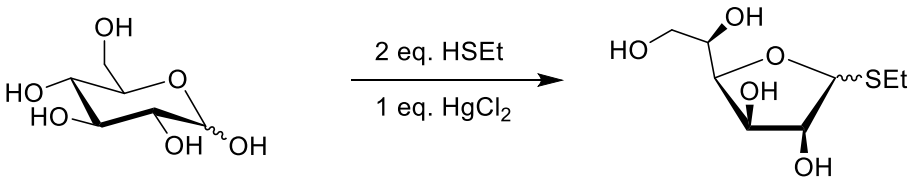
While DMC-mediated activation and substitution is effective for more acidic thiols such as aryl thiols, less acidic thiols show a tendency to form DMC adducts which interfere with the reaction efficiency (**Scheme 1.17**).[154] In order to avoid this side reaction, Novoa *et al*. developed a procedure based upon pre-activation of the oligosaccharide with DMC, followed

by addition of thiol once the sugar was fully activated. In this manner they carried out the DMC-mediated activation and nucleophilic substitution using the thiol of a cysteine residue of a peptide on a solid phase support as a nucleophile in order to prepare a glycopeptides. Moderate yields were observed for one reaction cycle, but by iteratively adding activated sugar they were able to obtain a quantitative yield for glycosylated peptide over three cycles.

DMC has a potential for modifying unprotected sugars that has just barely begun to be tapped. Further investigation into DMC could allow the direct preparation of all manner of unprotected carbohydrate derivatives, and from them a wide variety of glycoconjugates.

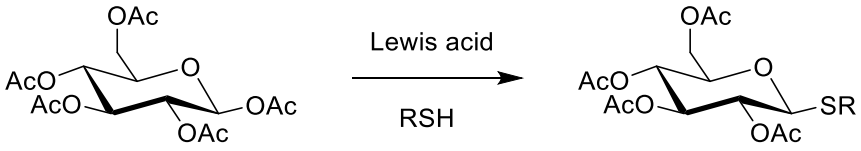
1.9 Thioglycosides

Sulfur is widely used in carbohydrate chemistry, most commonly in the form of thioglycosides. Thioglycosides were first prepared over 90 years ago by Schneider *et al*.[158] by the reaction of two equivalents of ethanethiol with glucose in the presence of a stoichiometric amount of mercuric chloride (**Scheme 1.18**).



Scheme 1.18: The preparation of thioglycosides by Schneider *et al*.

In modern times thioglycosides are typically prepared by the reaction of peracetylated sugars with a thiol in the presence of a Lewis acid (**Scheme 1.19**), most commonly $BF_3.OEt_2$,[159] but a number of other Lewis acids such as $SnCl_4$,[160] $ZrCl_4$,[161] and MoO_2Cl_2[162] have also been used.

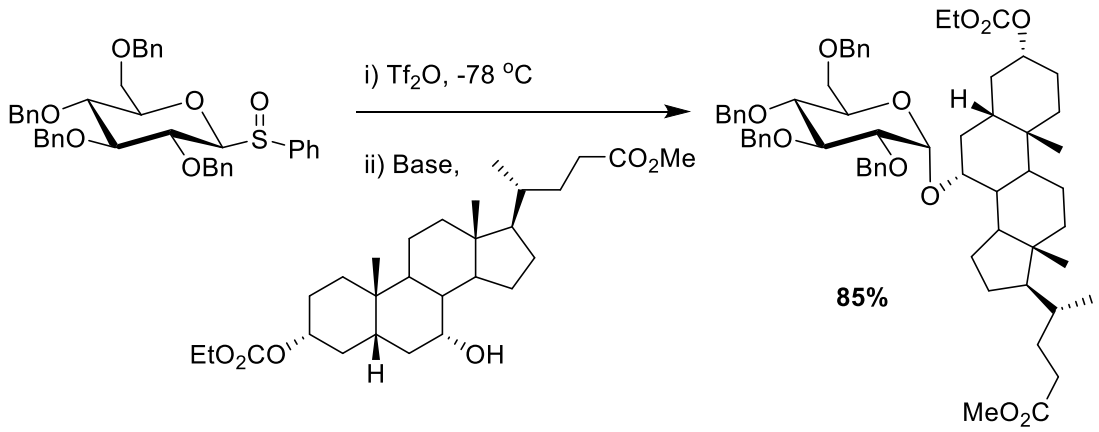


Scheme 1.19: Preparation of thioglycosides by the reaction of peracetylated monosaccharides with a Lewis acid and a thiol.

Thioglycosides are stable to a wide range of conditions, and are frequently used as glycosyl donors in glycosylation reactions, as they can be activated in a large number of ways, including with methyl triflate,[163, 164] *N*-iodosuccinimide together with a variety of other species,[165, 166] and various other conditions.[167, 168] However, in many cases they are less reactive than other glycosyl donors, such as trichloroacetimidates.[169]

1.10 Glycosyl sulfoxides

Although thioglycosides are only moderately reactive in glycosylation reactions, they are special in that they can be converted to glycosyl sulfoxides, which display significantly greater activity than the corresponding thioglycoside. Glycosyl sulfoxides were first used in glycosylation reactions by Kahne *et al.*[170] in 1989. They proved significantly more active than other glycosyl donors of the time. For example, the glycosylation of an extremely sterically demanding steroid unit (**Scheme 1.20**) was accomplished with a yield of 85%, while other attempts to synthesise similar species gave yields for the crucial glycosylation step of 40% or less.[171]



Scheme 1.20: Glycosylation of a a sterically demanding steroid unit by Kahne *et al.*[170]

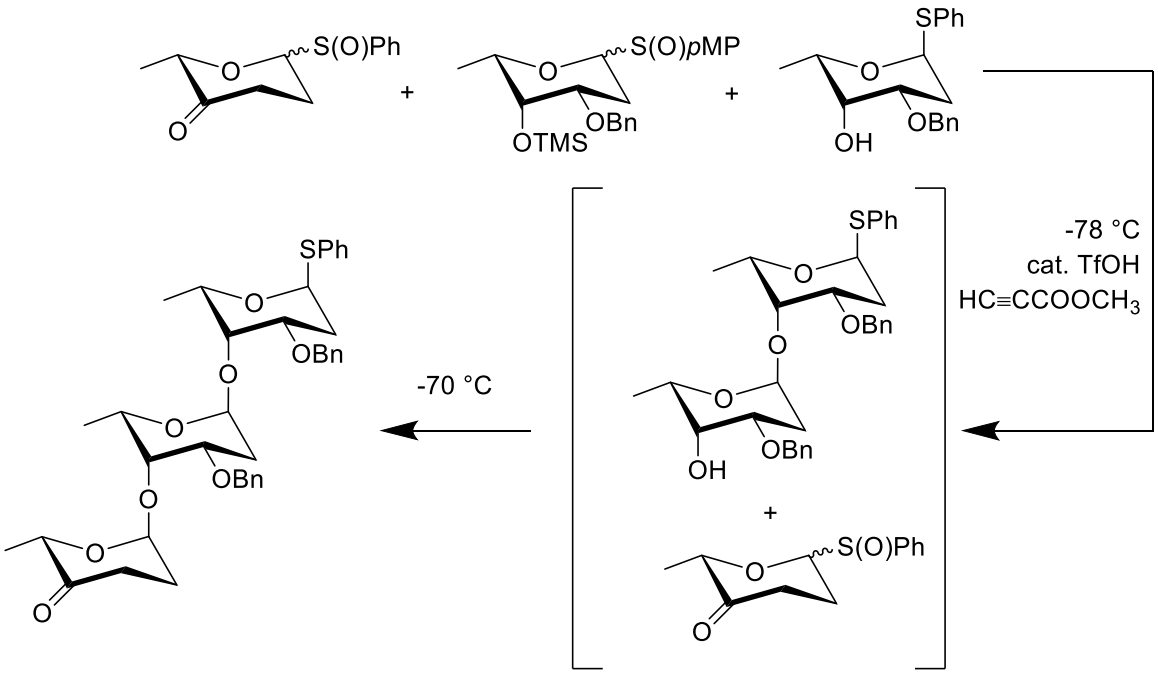
1.10.1 Applications of glycosyl sulfoxides

Since they were first reported glycosyl sulfoxides have been used in glycosylation reactions where their increased reactivity or other unique properties are important. Crich and co-workers have repeatedly used glycosyl sulfoxides in their syntheses of β-mannans.[172-175] In these syntheses the increased propensity of glycosyl sulfoxides to give the β-anomer when used in glycosylation reactions makes them ideal for the synthesis of chains of β-linked mannose units,

which are otherwise very difficult to access. This selectivity is proposed to be due to the formation of an α-glycosyl triflate when a glycosyl sulfoxide is activated by triflic anhydride,[176] which is subsequently selectively attacked in an S_N2 reaction to give a β-glycoside product.

Crich has also exploited both the increased reactivity and β selectivity of glycosyl sulfoxides by using them in the preparation of β-mannosyl phosphoisoprenoids.[172, 177] This class of molecule had in the past proved to be very difficult to prepare,[178] but using glycosyl sulfoxides as the glycosyl donors allowed their preparation in moderate yields with good selectivity, an improvement over other methods. Glycosyl sulfoxides are key to the preparation of these biologically relevant molecules.

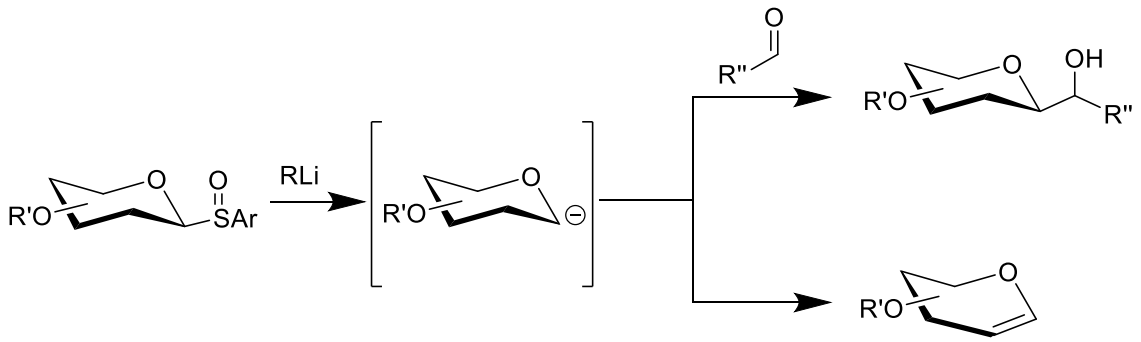
Like thioglycosides, glycosyl sulfoxides can be activated with a wide range of different activating agents. The first and most common activating agent used is triflic anhydride,[170] but reagents such as triflic acid[179] and trimethylsilyl triflate[180] have also been used. Significantly, because the reactivity of glycosyl sulfoxides is both much greater than that of thioglycosides and dependent on the aglycon unit of the sulfoxide, careful tuning of the reaction conditions allows selective reaction of sulfoxides both in the presence of thioglycosides[180, 181] and in the presence of other glycosyl sulfoxides.



Scheme 1.21: One pot synthesis of the trisaccharide component of cylcamycin by Kahne *et al*. [179]

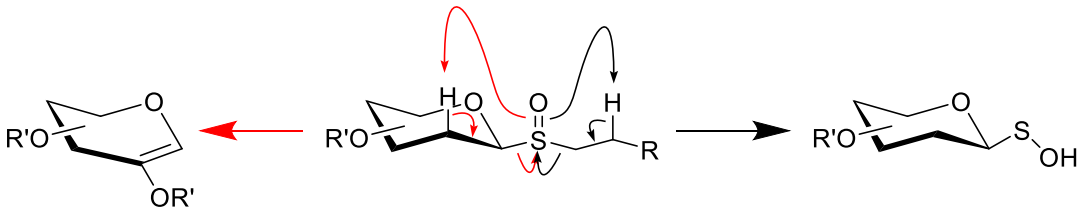
This is most elegantly demonstrated in the synthesis of the cyclamycin trisaccharide by Kahne *et al*.[179] (**Scheme 1.21**). In this synthesis three monosaccharide units were combined and activated in one pot to produce a trisaccharide with a thioglycoside at its reducing terminus. This was possible because the rate limiting step in glycosylation using glycosyl sulfoxides is triflation of the glycosyl sulfoxide, and therefore alteration of the electron donating ability of the aglycon, for example by adding an electron donating methoxy group to the phenyl ring, increases the rate of activation and hence the rate of glycosylation. This therefore allows selective activation of *p*-methoxyphenyl glycosyl sulfoxides in the presence of phenyl sulfoxides provided the conditions are carefully controlled. In addition, the use of a silyl protecting group which is stable at low temperatures (as used in the activation of *p*-methoxyphenyl glycosyl sulfoxides), but not at higher temperatures (as used in the activation of phenyl glycosyl sulfoxides) allows the *in situ* deprotection of a hydroxyl group in order to carry out the second glycosylation reaction selectively with respect to the glycosyl acceptor.

Furthermore, this synthesis could be carried out with a thioglycoside at the reducing terminus, allowing for further glycosylation if required, without any further modification of the structure. As can be seen from this synthesis, glycosyl sulfoxides are versatile reagents that can be subtly modified in order to afford significant control over a reaction.



Scheme 1.22: Use of glycosyl sulfoxides to generate glycosyl carbanions and subsequent reactions of glycosyl carbanions.

In addition to glycosylation reactions, glycosyl sulfoxides can be used in a number of other reactions. Treatment of glycosyl sulfoxides with an alkyl lithium such as *n*-butyl lithium or *t*-butyl lithium generates a glycosyl carbanion which can then either react with an electrophile, for example to synthesise a *C*-glycoside,[182] or eliminate to produce a glycal[183] (**Scheme 1.22**).



Scheme 1.23: Thermal rearrangement of glycosyl sulfoxides.

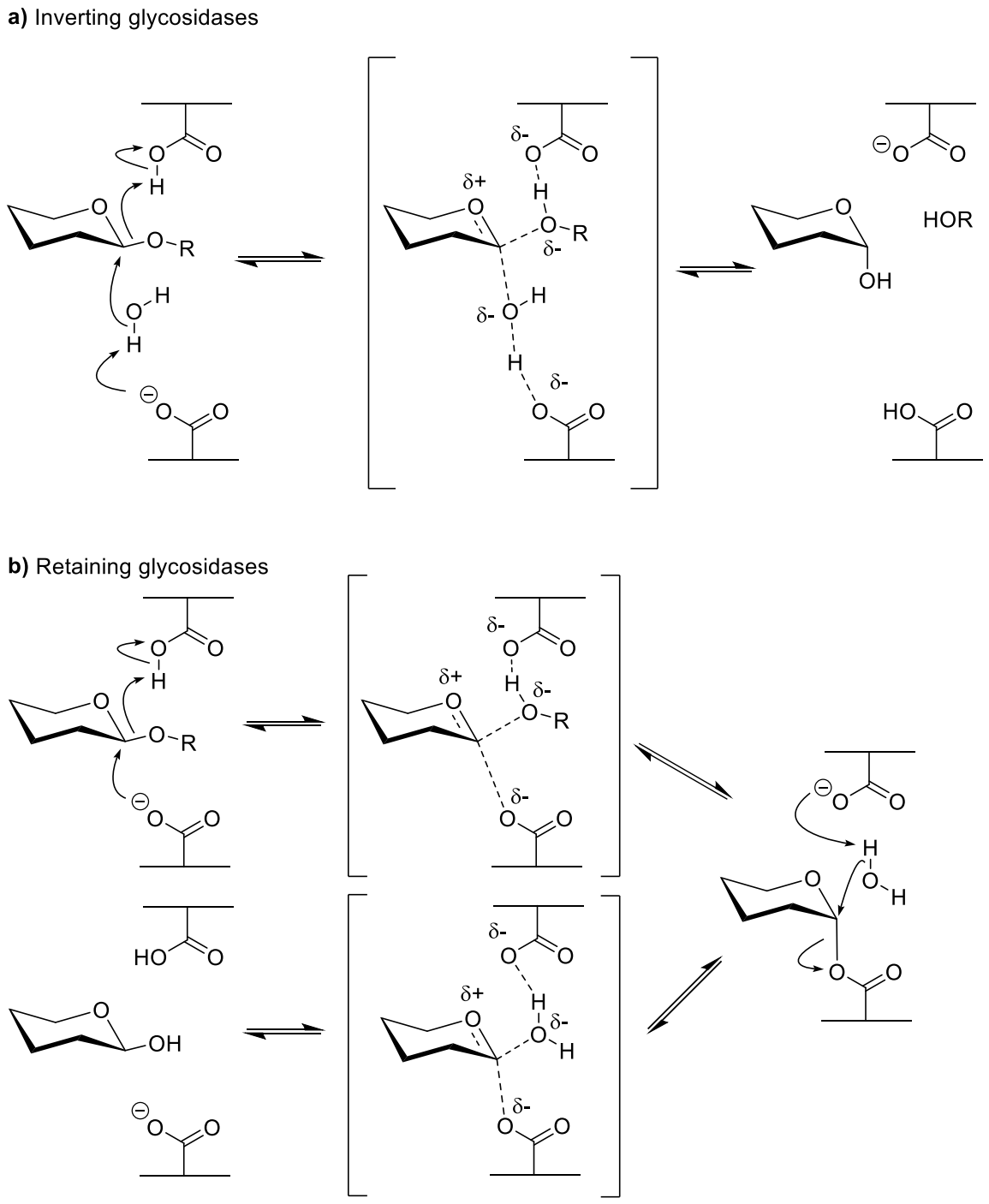
Upon heating, sulfoxides undergo β-elimination to form a sulfenic acid and an alkene (**Scheme 1.23**). Depending on which direction β-elimination occurs in this can produce two different classes of products. Typically glycosyl sulfoxides will β-eliminate to form 2-hydroxy or 2-amino glycals, which are useful synthetic intermediates.[184] If aromatic glycosyl sulfoxides are

used this is the only potential elimination pathway. Alkyl glycosyl sulfoxides, however, may also β-eliminate towards the aglycon, producing glycosyl sulfenic acids. This process has been studied in depth by Aversa *et al*.,[185-187] who determined that through careful choice of the aglycon group glycosyl sulfenic acids could be generated by heating glycosyl sulfoxides, and subsequently be reacted *in situ* with either alkynes to generate α,β-unsaturated sulfoxides, or with thiols to generate mixed disulfides.

Sulfoxides have also found use as glycosidase inhibitors. Glycosidases are an extremely widespread category of enzymes which catalyse the hydrolysis of glycosidic linkages.[188, 189] As glycosidases are involved in such a wide variety of processes, inhibitors of glycosidases have a correspondingly wide variety of applications, including glycogen phosphorylase inhibitors for the treatment of type 2 diabetes,[190] neuraminidase inhibitors for the treatment of influenza,[191] inhibitors of glycosphingolipid formation for the treatment of Gaucher's disease,[192] and many more.[193]

1.11 Glycosidase mechanisms

Most glycosidase enzymes share one of two general mechanisms (**Scheme 1.24**).[194] Inverting glycosidases rely upon acid-catalysed activation of the glycoside to be cleaved while simultaneously activating water for attack by base catalysis. This results in a one-step mechanism whereby the stereochemistry at the anomeric centre is inverted. The transition state of this mechanism involves the build-up of positive charge on the endocyclic sugar oxygen, and protonation of the exocyclic sugar oxygen, followed by a build-up of negative charge.[195] A water molecule is also present in the active site of the enzyme.



Scheme 1.24: General mechanism for inverting (**a)**) and retaining (**b)**) glycosidases.

Retaining glycosidases display a similar active site to inverting glycosidases, but carry out the hydrolysis reaction slightly differently. The glycoside is again activated by protonation of the exocyclic oxygen, however in this case the nucleophile is not water but a deprotonated carboxylic acid residue. A covalent enzyme-carbohydrate intermediate is then formed. The

former catalytic acid is now deprotonated and can therefore act as a catalytic base to activate water for attack and hydrolysis of the covalent intermediate to form the hydrolysed product. Due to the double inversion of stereochemistry, the reactant and product share the same stereochemistry at the anomeric centre.

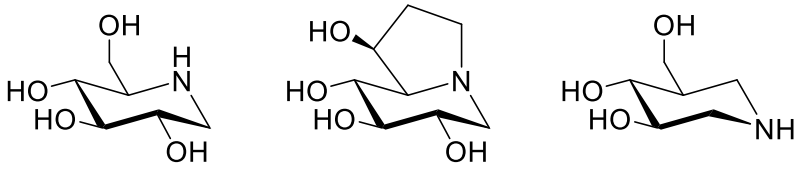
The first transition state for retaining glycosidases is similar to that of inverting glycosidases, in that there is again protonation of the exocyclic oxygen and a build-up of positive charge on the endocyclic oxygen. However, there is no water molecule present in the active site of the enzyme, with a covalent intermediate being formed instead. This naturally leads to a second transition state, in which the covalent intermediate is hydrolysed by a water molecule to give hydrolysed product. This process is base-catalysed, as the catalytic residue assists in deprotonating the water so it can attack the anomeric centre of the sugar. Again, there is a significant build-up of positive charge on the endocyclic oxygen.

Binding of the carbohydrate substrate in the active site also results in significant distortion of the carbohydrate structure.[196] Crystallographic studies have shown that the ring is distorted from a standard $^{4}C_{1}$ chair conformation to a $^{1}S_{3}$ skew-boat conformation.[197] Leaving group (aglycon) binding is essential for this to take place, as species which do not bind in the leaving group pocket have been shown to be undistorted.

1.12 Glycosidase inhibitors

The concept that enzymes have the highest affinity for the transition state of a reaction in order to lower the energy of activation of the reaction is well established.[198] One possible logical approach to inhibitor design is therefore to design species which mimic the transition state of the reaction, as this is the species to which the enzyme binds most strongly. Most glycosidase inhibitors reflect this design philosophy, although in a number of cases detailed studies have indicated that binding is at most partially due to the compound in question being a transition state mimic. Other interactions both within and external to the enzyme active site can result in inhibitors which bind strongly to an enzyme, and in many cases this is true for glycosidase inhibitors, which often form strong salt bridge interactions with catalytic residues.

1.12.1 Inhibitors containing endocyclic heteroatoms



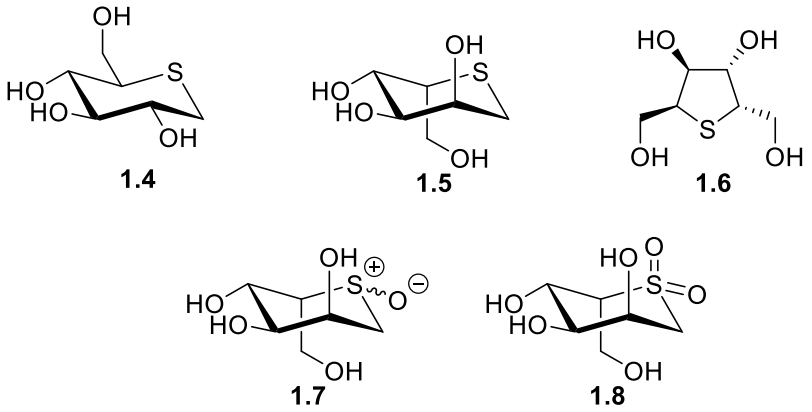
1.1 **1.2** **1.3**

Figure 1.8: Amine-based glucosidase inhibitors.

As noted previously there is significant build-up of positive charge on the endocyclic oxygen, and so mimicking this feature might be expected to be a step towards more potent inhibitors. Commonly, a positively charged heteroatom is introduced in place of the endocyclic oxygen or the anomeric carbon, a feature which could be expected to mimic this partial positive charge. Furthermore, glycosidase active sites contain multiple carboxylate residues, and charged species would be expected to form strong interactions with these residues. This can be seen in potent glycosidase inhibitors such as 1-deoxynojirimycin **1.1**[199, 200] and castanospermine **1.2**,[201]

both glucosidase inhibitors (**Figure 1.8**). Protonation of the ring nitrogen, which occurs fully at physiological pH, creates a positive charge where the endocyclic oxygen would be, increasing the binding affinity of the molecule for the glycosidase. Isofagomine **1.3** was developed based on a similar principle, but with a nitrogen atom in place of the anomeric carbon, and is also a strong inhibitor of glucosidases.[202]

Nitrogen is not the only atom that has been used in this role. Sulfur has been investigated in glycosidase inhibitors, both as uncharged sulfides and charged sulfonium salts, sulfoxides, and sulfones.



Compound	Inhibition of α-glucosidase from *Bacillus stearothermophilus* at 1 mM inhibitor concentration
1.4	6%
1.5	18%
1.6	18%
1.7	22%
1.8	15%

Figure 1.9: 5-Thiosugars and derivatives prepared by Le Merrer *et al*. and their inhibition of α-glucosidase from *Bacillus stearothermophilus*.[203]

Sulfides are not expected to be good glycosidase inhibitors, as they will generally share both the geometry and non-polarised bonds of the starting material, rather than the transition state.

They are also uncharged, and as a result do not interact strongly with the catalytic carboxylate residues. This is borne out by experiment, whereby the preparation and inhibition testing of 5-thio analogues of 1-deoxynojirimycin and similar structures (**Figure 1.9**) by Le Merrer *et al*.[203] revealed that thioether compounds **1.4**, **1.5**, and **1.6** displayed only modest inhibition against a panel of glycosidase enzymes, with minimal inhibition even at a concentration of 1 mM. In this case, a limited study of the effect of oxidation of the sulfur demonstrated a limited effect on the inhibition of the compound in question, with both sulfoxide **1.7** and sulfone **1.8** displaying similar inhibition percentages. Sulfoxide **1.7** did demonstrate slightly improved inhibition, but the increase was minimal.

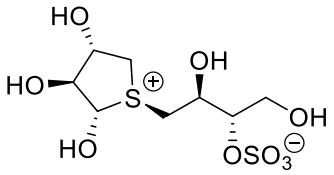


Figure 1.10: The structure of salacinol

Salacinol (**Figure 1.10**) is a natural product derived from the roots and stems of *Salacia reticulate*.[204] It is a strong glucosidase inhibitor, displaying comparable inhibition of rat intestinal maltase and sucrase to acarbose, a leading glucosidase inhibitor used for the treatment of type-2 diabetes.[205] In addition, salacinol displayed greatly increased inhibition of isomaltase over acarbose. The presence of a positively charged sulfonium ion is presumably a significant factor in salacinol's strong binding to glucosidases.

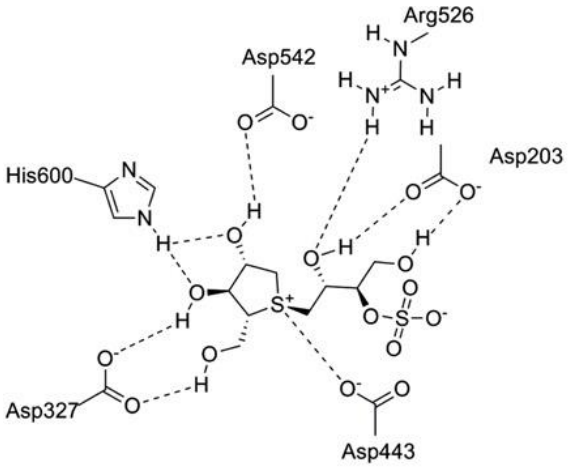


Figure 1.11: Binding of salacinol to the active site of NtMGAM. Adapted from the work of Nakamura *et al*.[206]

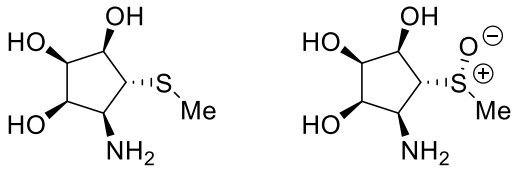
Salacinol has since been synthesised[207] and its docking to the *N*-terminal catalytic domain of maltase-glucoamylase (NtMGAM) has been studied (**Figure 1.11**).[206] As would be expected, the hydroxyl groups form a network of hydrogen bonds, presumably mimicking interactions formed by the natural substrate. The sulfonium ion forms a salt bridge with an aspartate residue, a feature consistent with iminosugar inhibitors. Investigation has suggested that perhaps this interaction is less representative of transition state mimicry and more an example of fortuitous binding, whereby the inhibitor takes advantage of different interactions to the natural substrate, and therefore is not a transition state mimic.[208] This is logical, as while there is a build-up of positive charge in the transition state, it is only partial, not a full charge as is observed in sulfonium or iminosugars. While fortuitous binders are no less effective for not being transition state mimics, they are typically harder to discover in a logical manner, and modification will not necessarily result in predictable variation in binding affinity.

Strong hydrogen bonding interactions or salt bridges are characteristic of glycosidase inhibitors that are not transition state mimics.[208] Multiple carboxylate residues are invariably present in

the active site of glycosidase enzymes to act as catalytic acids, bases, or nucleophiles, and therefore charged compounds that depend on interactions with these residues often show inhibition of a wide variety of glycosidases, despite not being transition state mimics.

1.12.2 Inhibitors containing exocyclic heteroatoms

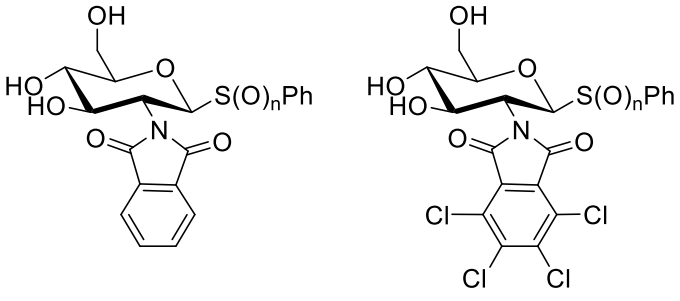
As the exocyclic oxygen is protonated during the glycosidase-catalysed hydrolysis of glycosides, the exocyclic oxygen may also develop a partial positive charge, depending on the time of protonation on the reaction co-ordinate relative to carbon-oxygen bond cleavage. In addition, a positive charge can help create salt bridge interactions within the enzyme active site, as discussed previously. Molecules bearing a positive charge or partial positive charge in the position of the exocyclic oxygen have therefore also been investigated.



1.9 **1.10**

Figure 1.12: The structure of Mannostatins A and B.

For example, Mannostatin A (**1.9**, **Figure 1.12**) and its sulfoxide Mannostatin B (**1.10**) have been shown to be strong inhibitors of rat epididymis mannosidase, both displaying K_is of 48 nM.[209] In this case modelling has suggested that the significant interaction is hydrogen bonding of the amine group with the general acid/base aspartic acid residue.[210] Again, the mannostatins do not appear to be true transition state mimics, instead inhibiting due to strong hydrogen bonding interactions. However, the five-membered ring is proposed to be a good mimic for the distorted six-membered ring of the transition state.



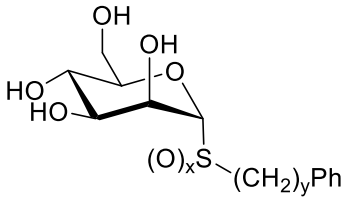
1.11: n = 0
1.12: n = 1

1.13: n = 0
1.14: n = 1

Compound	Inhibition of β-glucosidase from *Caldocellum saccharolyticum* at 1 mM inhibitor concentration
1.11	45%
1.12	69%
1.13	0%
1.14	82%

Figure 1.13: Inhibitors developed by Robina *et al*. and their inhibition of β-glucosidase from *Caldocellum saccharolyticum*.[211]

Glycosyl sulfoxides have also been investigated as glycosidase inhibitors featuring a polarised exocyclic group. Robina *et al*. investigated 2-phthalamido derivatives **1.11**-**1.14** as inhibitors of β-glucosidases (**Figure 1.13**).[211] Thioglycoside **1.11**, its sulfoxide **1.12**, and chlorinated sulfoxide **1.14** displayed moderate inhibition at a 1 mM concentration, with chlorinated sulfoxide **1.14** being the most effective inhibitor. Curiously, **1.14** inhibits uncompetitively at substrate concentrations below 1 mM. Uncompetitive inhibition is when the inhibitor only binds to the enzyme substrate complex, implying that **1.14** does not occupy the sugar's position in the enzyme active site. It is possible that **1.14** is binding in another location in the active site, or is binding elsewhere in the molecule.



1.15: x = 0, y = 1
1.16: x = 0, y = 2
1.17: x = 1, y = 1
1.18: x = 1, y = 2
1.19: x = 2, y = 1
1.20: x = 2, y = 2

Compound	IC_{50} for dGMIIb (mM)
1.15	No inhibition
1.16	No inhibition
1.17	No inhibition
1.18	2.5
1.19	2.0
1.20	1.5

Figure 1.14: Inhibitors of dGMIIb developed by Poláková *et al*.[212]

Mannose thioglycosides, sulfoxides, and sulfones **1.15-1.20** were investigated by Poláková *et al*. as inhibitors of golgi α-mannosidase II (dGMIIb).[212] Compounds **1.18**, **1.19**, and **1.20** displayed moderate inhibition of dGMIIb, with IC_{50} values between 1.5 and 2.5 mM (**Figure 1.14**). The other compounds did not inhibit dGMIIb, indicating that inhibitory activity was enhanced by oxidation of the sulfur atom, and also by increasing the flexibility of the alkyl chain. Docking studies showed that a hydrogen bond between the catalytic acid residue of the mannosidase and an oxygen of the sulfoxide or sulfone plays a role in the binding of the inhibitor to the mannosidase, similarly to iminosugars.

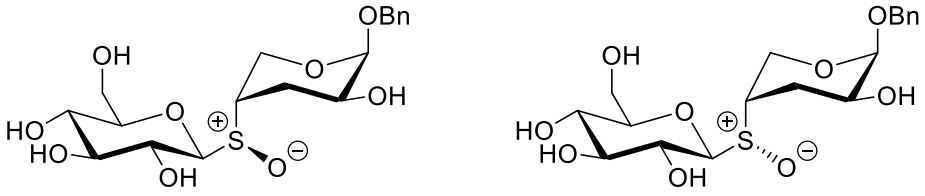


Figure 1.15: Thiodisaccharide sulfoxides investigated by Colomer *et al*.[213]

Colomer *et al*. investigated thiodisaccharide sulfoxide diastereoisomers **1.21 *R*** and **1.21 *S*** (**Figure 1.15**) as inhibitors of β-galactosidase from *E. coli*.[213] Both displayed moderate K_is of 0.45 mM and 0.19 mM respectively. Interestingly, **1.21 *S*** was also shown to be a substrate for the glycosidase.

While species such as sulfoxides do not normally constitute transition state mimics, they have been shown to be able to interact with catalytic residues in the active site of glycosidase enzymes. These interactions are another potential route to the inhibition of glycosidases, as shown by the widespread use of iminosugars and similar structures as glycosidase inhibitors.[214]

1.13 Synthesis of glycosyl sulfoxides

Glycosyl sulfoxides show wide utility in chemical reactions, both in their traditional role as glycosyl donors and in other sequences, and in addition show some promise as glycosidase inhibitors. However, their use is limited by the difficulty of preparing them. The most common procedure for preparing glycosyl sulfoxides is Kahne's original procedure, oxidation of the corresponding thioglycoside with *m*CPBA.[170] *m*CPBA is relatively inexpensive and safe to work with, and is therefore widely used as an organic oxidising agent, but the use of *m*CPBA in the synthesis of glycosyl sulfoxides has some notable drawbacks.[215] Firstly, careful control

of the oxidation conditions is required to avoid over-oxidation of the thioglycoside to the unreactive glycosyl sulfone. This typically means the reaction must be carried out at low temperatures, and with a carefully controlled concentration of *m*CPBA. In addition to the reduced yield resulting from over- or under-oxidation, glycosyl sulfoxides can be difficult to separate from both the starting thioglycoside and undesired sulfone due to their similar polarities, and therefore even a small amount of impurity can be a significant issue, particularly when compared to glycosyl donors such as trichloroacetimidates,[169] which display comparable activity to glycosyl sulfoxides in most cases. In addition, the use of *m*CPBA as an oxidant results in the formation of a significant amount of *m*-chlorobenzoic acid as a byproduct, which can again be difficult to separate from the glycosyl sulfoxide. Trichloroacetimidates and other glycosyl donors, including thioglycosides, are typically preferred over glycosyl sulfoxides even for difficult glycosylation reactions as the increased efficiency of glycosylation when using glycosyl sulfoxides is countered by the increased difficulty of preparing them. The use of glycosyl sulfoxides in glycosylation reactions is in large part restricted to cases where other glycosyl donors have failed, and other reactions in which glycosyl sulfoxides are used typically find limited application.

Since the first use of glycosyl sulfoxides, a number of attempts have been made to develop alternative oxidation methods. The most widely used was developed by Kakarla *et al*.[215] and relies upon the ability of silica gel to act as an adsorbent for peroxy oxidising agents, increasing the rate at which they react.[216] The treatment of thioglycosides with hydrogen peroxide, acetic anhydride, and silica gel in dichloromethane resulted in the oxidation of sulfides to sulfoxides, while producing little to no sulfone. This method avoids a number of the drawbacks associated with the use of *m*CPBA, as significantly less care is required to avoid over-oxidation: the reaction can be carried out at room temperature, and the reagent system demonstrates a reduced

propensity to overoxidise to the sulfone. Separation of the reagents and byproducts is also significantly easier, as all reagents used may be filtered off or are water soluble. The most significant drawback of this method is that the reagent system used is significantly less active than *m*CPBA: extended reaction times are required, and less reactive thioglycosides typically display a low conversion to glycosyl sulfoxide. *m*CPBA reacts significantly more quickly, and will oxidise a wider variety of thioglycosides efficiently.

These two methods have both been modified in order to attempt to overcome their limitations. Agnihotri *et al.*[217] demonstrated that the use of potassium fluoride as an additive to the *m*CPBA oxidation of thioglycosides results in complete selectivity for sulfoxide, avoiding any oxidation to sulfone. Requiring limited control over reaction conditions and causing very rapid oxidation to sulfoxide (<10 minutes), this procedure is a significant improvement over unmodified *m*CPBA oxidation. However, *m*-chlorobenzoic acid will still be produced, and therefore purification of the products will still be challenging.

An attempt was made by Chen *et al*. to modify Kakarla's $H_2O_2/Ac_2O/SiO_2$ procedure by substituting other peroxy oxidising agents for hydrogen peroxide.[218] *tert*-Butyl hydroperoxide and oxone® were both tested, and both proved to be less reactive than hydrogen peroxide by a significant margin. Extended reaction times are therefore required (>20 h in almost all cases, and up to 42 h) even for relatively electron-rich, and therefore easily oxidised, thioglycosides. The conditions may be useful for selective oxidations when more than one thioether is present, and avoid the use of hydrogen peroxide, which can be dangerous, but otherwise are inferior to Kakarla's original conditions.

A number of unrelated methods have also been reported for the oxidation of thioglycosides to glycosyl sulfoxides. Fluorinating reagents, including Selectfluor®[219] and 1-fluoropyridinium triflates,[220] produce glycosyl sulfoxides in very high yields by first fluorinating the sulfur of the thioglycoside, forming a sulfur-fluorine species which is subsequently hydrolysed to form a glycosyl sulfoxide. Selectfluor® produced glycosyl sulfoxides for a number of thioglycosides cleanly and in high yields, but it is noted the reaction produces hydrofluoric acid, and strong acids can often cause the decomposition of larger oligosaccharides. In addition, the fluorinating reagents themselves are highly reactive, and could potentially damage less robust structures. The use of fluorinating reagents for the oxidation of larger oligosaccharides is therefore likely to be limited.

Chen *et al*. explored the use of microwave irradiation to accelerate the oxidation of thioglycosides.[221] Investigation revealed that magnesium monoperoxyphthalate was the most effective oxidant under these conditions, proving slightly more effective than sodium periodate adsorbed on silica gel and oxone®. The conditions developed resulted in significantly more rapid reaction than Kakarla's $H_2O_2/Ac_2O/SiO_2$ conditions, with oxidation taking less than 1 h even for less reactive thioglycosides. However, microwave irradiation is not widely used, and while selectivity for the sulfoxide was observed any excess magnesium monoperoxyphthalate caused the formation of sulfone, which means that careful control is required to avoid over-oxidation.

Huang *et al.* examined the use of porphyrin catalysts and sodium hypochlorite as oxidant.[222] Manganese-porphyrin complexes proved to be moderately effective as catalysts, providing a number of thioglycosides in good yields. However, the yields were lower than those obtainable

by other methods, the more effective catalysts are not readily commercially available, and the reaction must be carried out under phase-transfer conditions, which means that the reaction is less effective for less polar sugars. However, the method is highly selective for the production of single sulfoxide diastereoisomers as compared to other oxidation methods, which is useful when a single sulfoxide diastereoisomer is required, for example for use as a glycosidase inhibitor.

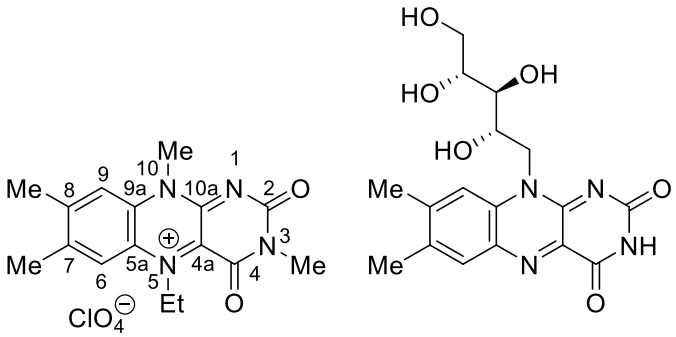
Very recently Wang *et al*. developed a method based on hydrogen peroxide and phenol,[223] either using phenol as the solvent or dissolved in acetic acid. These conditions were effective for more reactive substrates, but suffered from low conversions when less reactive substrates (such as nitrophenyl thioglycosides) were used. In addition, the highly acidic conditions proved detrimental when thioglycosides with acid-labile protecting groups, such as acetals and silyl protecting groups, were used as substrates, generally resulting in low yields of the corresponding glycosyl sulfoxide. Also, separation of phenol from the product of the reaction was reportedly difficult.

A number of varied techniques for the oxidation of thioglycosides to glycosyl sulfoxides have been explored, but at this time none have succeeded in all areas. No one method provides an efficient reaction with simple purification and no over-oxidation to sulfone. Glycosyl sulfoxides therefore find only limited application, as often other species will fill similar roles more efficiently or with less complexity (for example, trichloroacetimidates or even thioglycosides for glycosylation reactions). A simple, clean method for the synthesis of glycosyl sulfoxides would allow more widespread utilisation of reactions they are involved in, and make the development of new applications more attractive.

1.14 Methods of sulfide oxidation to sulfoxide

The oxidation of sulfides to sulfoxides has been very well explored. Metal catalysed reactions are very widely used,[224-227] and although they are undoubtedly effective, often metal catalysts are expensive, unsustainable, and difficult to work with. There are also a number of reports of oxidations catalysed by entirely organic reagents and simple oxidising reagents.[228-231] Ideally the conditions applied to the oxidation of thioglycosides would be compatible with a range of conditions and functional groups, oxidise thioglycosides rapidly and selectively to the sulfoxide, use a simple to acquire and work with catalyst, and result in little or no byproducts. Hydrogen peroxide or oxygen would be preferred as an oxidant, as both are readily available, easy to work with, and produce only water as a byproduct.

1.14.1 Flavinium salt-catalysed oxidations

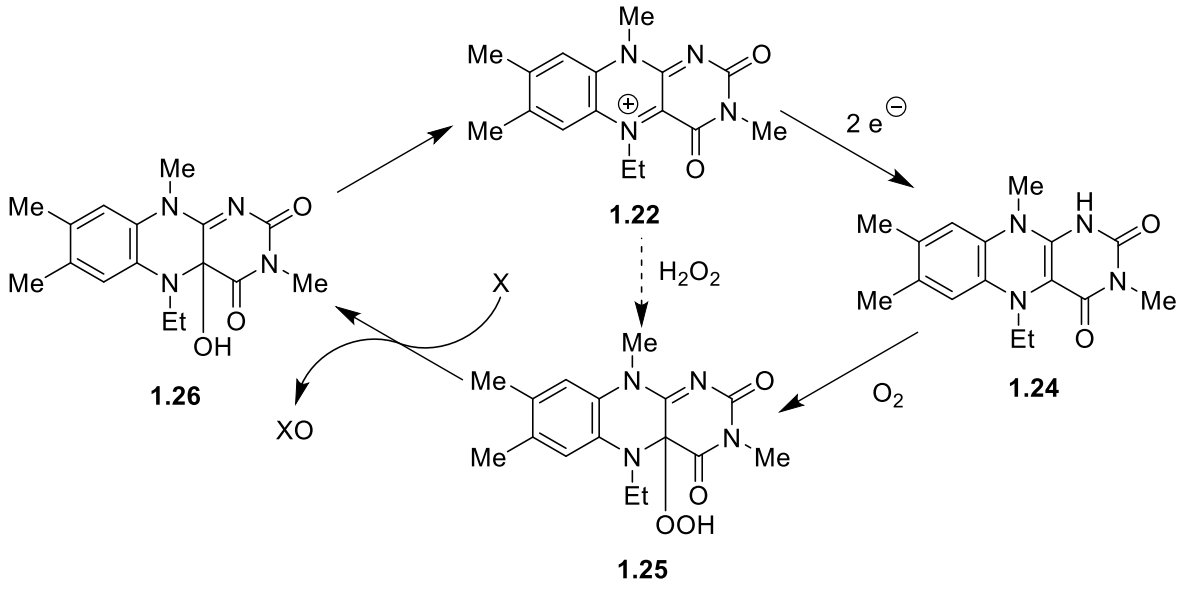


1.22 **1.23**

Figure 1.16: Flavinium salt catalyst used by Murahashi (**1.22**)[232] and riboflavin (**1.23**). Flavin atom numbering scheme shown on **1.22**.

In 1989 Murahashi reported the use of a flavinium salt species (**1.22**, **Figure 1.16**) to catalyse the oxidation of sulfur and nitrogen-containing molecules.[232] Flavins are found in nature as cofactors for oxidase, oxygenase, and dehydrogenase enzymes, normally as flavin adenine

dinucleotide (FAD) or flavin mononucleotide (FMN),[233] and this is therefore an example of biomimetic catalysis. Both FAD and FMN are biosynthesised from riboflavin (vitamin B2, **1.23**),[234] and all three biomolecules, as well as Murahashi's catalyst, contain a similar core structure, which is responsible for the catalytic activity. As would be expected, they also rely upon similar reactive intermediates in order to carry out catalysis. A catalytic cycle for **1.22** is shown in **Scheme 1.25**.[235] The 4a-hydroperoxy adduct **1.25** is believed to be the reactive intermediate responsible for the key oxidation step, as upon isolation and use in a non-catalytic reaction **1.25** was demonstrated to react more than 4 orders of magnitude faster than hydrogen peroxide with amines to convert them to their *N*-oxides.[236] Note that the labelling of atoms in all flavins and related species is displayed in **Figure 1.16**.



Scheme 1.25: Catalytic cycle for flavin oxidation catalysts.

A notable feature of this cycle is that the precise mechanism differs depending on the oxidising species. In nature, molecular oxygen is the most readily available species for regenerating **1.25**. In oxygenases, in order to generate reactive intermediate **1.25** using molecular oxygen the oxidised form of flavinium salt **1.22** is first reduced to species **1.24**, typically by NADPH,[237]

which then undergoes a single electron transfer process to form a biradical complex.[238] The flavin and superoxide radicals then combine to produce **1.25**.

This sequence of reactions can be bypassed if hydrogen peroxide is used to regenerate **1.25**. Peroxide can directly attack flavinium salt **1.22** at position 4a to form **1.25**, without the need for electron transfer from a reducing agent.[232] In this case the rate limiting step in this catalytic cycle is believed to be loss of water from species **1.26** to regenerate catalyst **1.22**.[232]

Since the first report on the use of flavinium salt **1.22** as a catalyst a number of groups have investigated the use of flavinium salts as catalysts for oxidation reactions. In addition to the amine and sulfide substrates investigated by Murahashi,[232] flavinium salts have since also been used to hydroxylate aromatics,[239] oxidatively hydroxylate boronic acids,[240] and carry out Baeyer-Villiger oxidations.[241, 242] However, amines and sulfides remain the most common substrates.

The flavinium salt catalysts in use today for oxidations fall into three broad categories based on their nitrogen substitution pattern. All three have alkyl substituents on two of N^1, N^5 and N^{10}, and this has been shown to be necessary for activity.[241]

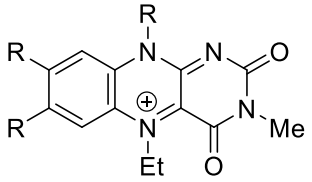
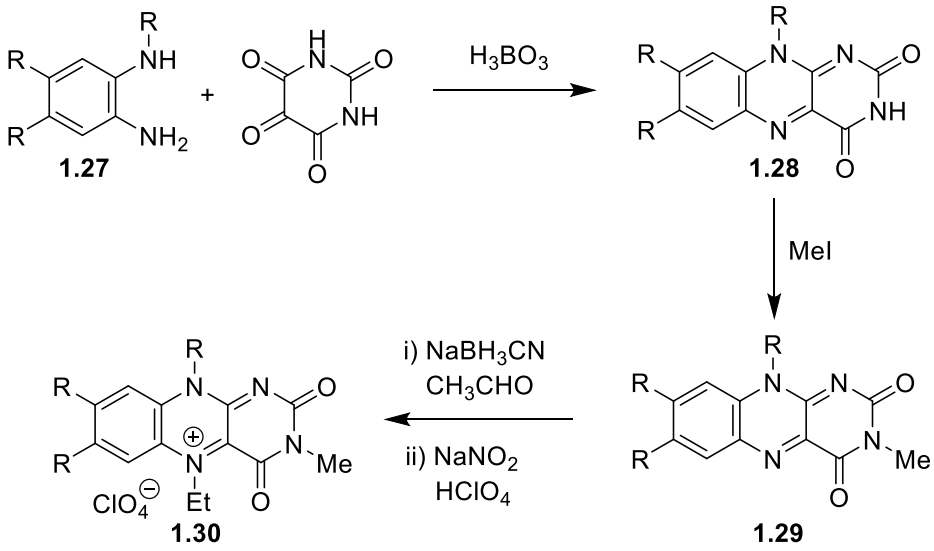


Figure 1.17: N^5,N^{10} substituted flavinium salts.

The first type to be developed were the N^5,N^{10} substituted flavinium salts (**Figure 1.17**).[232] This class of catalysts is the most similar to FAD of the three, with an extra substituent (typically an ethyl group) on N^5. N^3 is also generally methylated.

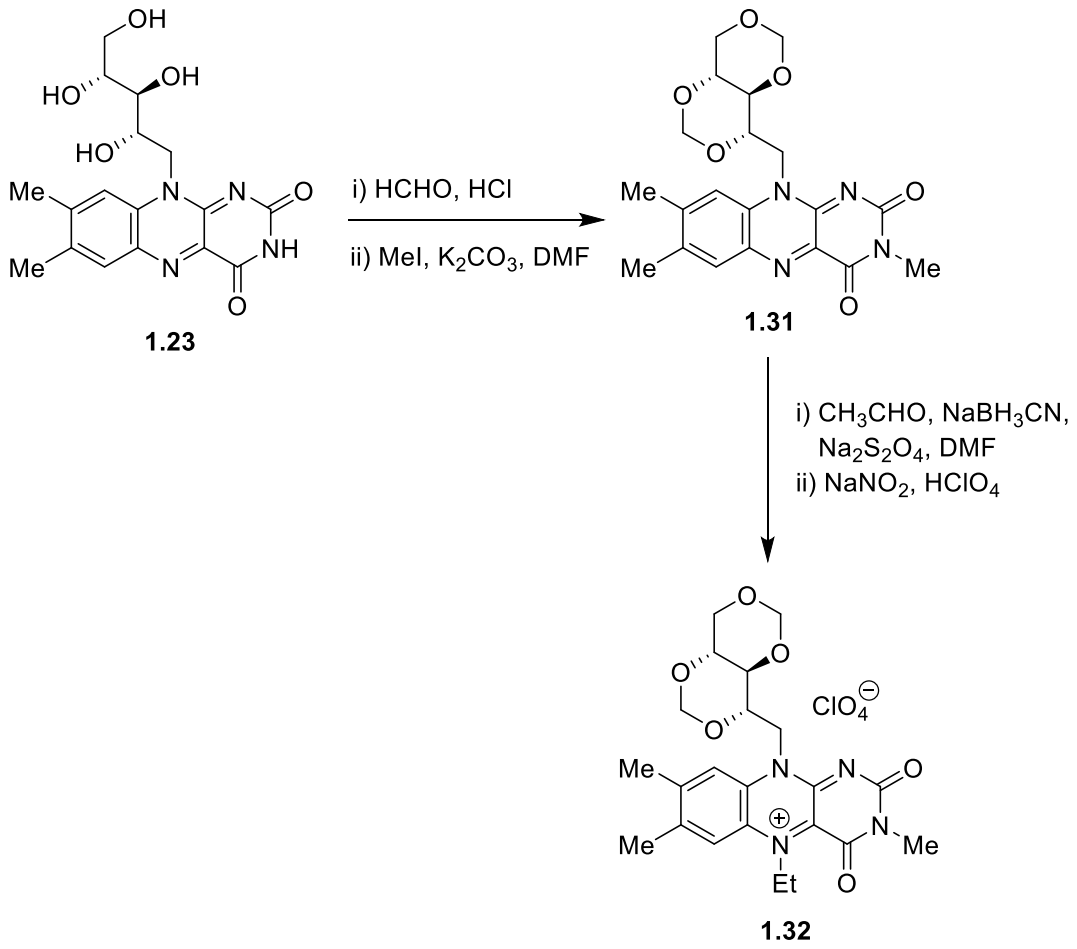
The standard method of preparing N^5,N^{10} flavinium salts is condensation of alloxan with a 1,2-diamino benzene derivative **1.27** to give flavin **1.28**, followed by methylation of N^3 with methyl iodide to produce substituted flavin **1.29**, and then reductive amination of N^{10} and salt exchange to afford flavinium salt **1.30** (**Scheme 1.26**).[243]



Scheme 1.26: Preparation of N^5,N^{10} substituted flavinium salts.

This method can be modified by altering the substituents on the 1,2-diaminobenzene or the aldehyde used for the reductive amination, allowing both subtle and significant modification to the structure of the catalyst. There are limited systematic comparisons of the activity of differentially substituted N^5,N^{10} substituted flavinium salts. N^5 is invariably ethyl substituted, but a variety of substituents for N^{10} have been examined, including short[232] and long[244] alkyl chains, aromatic groups,[241, 245] chiral groups,[246] and dimeric structures.[247, 248]

Catalyst **1.32** has been prepared by Imada *et al*.[242] (**Scheme 1.27**). Flavinium salt **1.32** is significant because instead of constructing the flavin skeleton from discrete building blocks, its synthesis instead begins with riboflavin **1.23**, the precursor to FAD. While this does make modification of the structure of **1.32** more difficult, the use of a widely available biomolecule as a starting material means that the majority of the required structure is already in place, and potentially makes **1.32** or similar structures less costly to produce on a large scale.

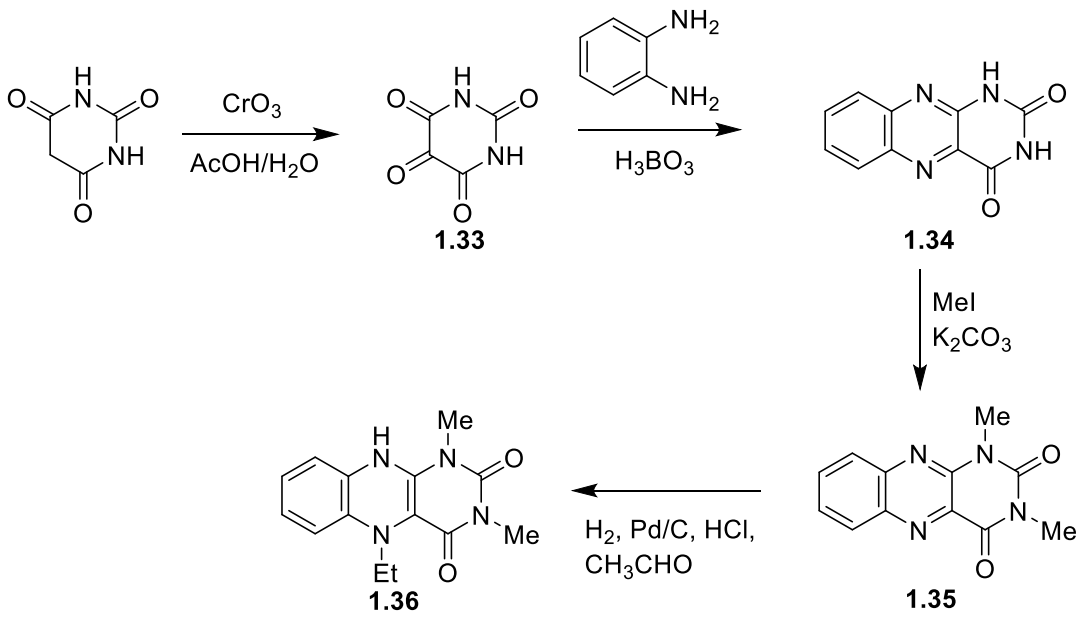


Scheme 1.27: Preparation of catalyst **1.32** from riboflavin by Imada *et al*.[242]

Riboflavin **1.23** is first converted to the diacetal under acidic conditions and methylated on N^3 to afford intermediate **1.31**. Then, as previously seen in **Scheme 1.26**, a reductive amination is performed to ethylate N^5, and the counterion is exchanged for perchlorate to produce flavinium salt catalyst **1.32**. Catalyst **1.32** has been shown to be active in the hydrogen peroxide oxidation

of sulfoxides, producing sulfoxides selectively and in high yields, similarly to other N^5,N^{10} substituted flavinium salt catalysts.

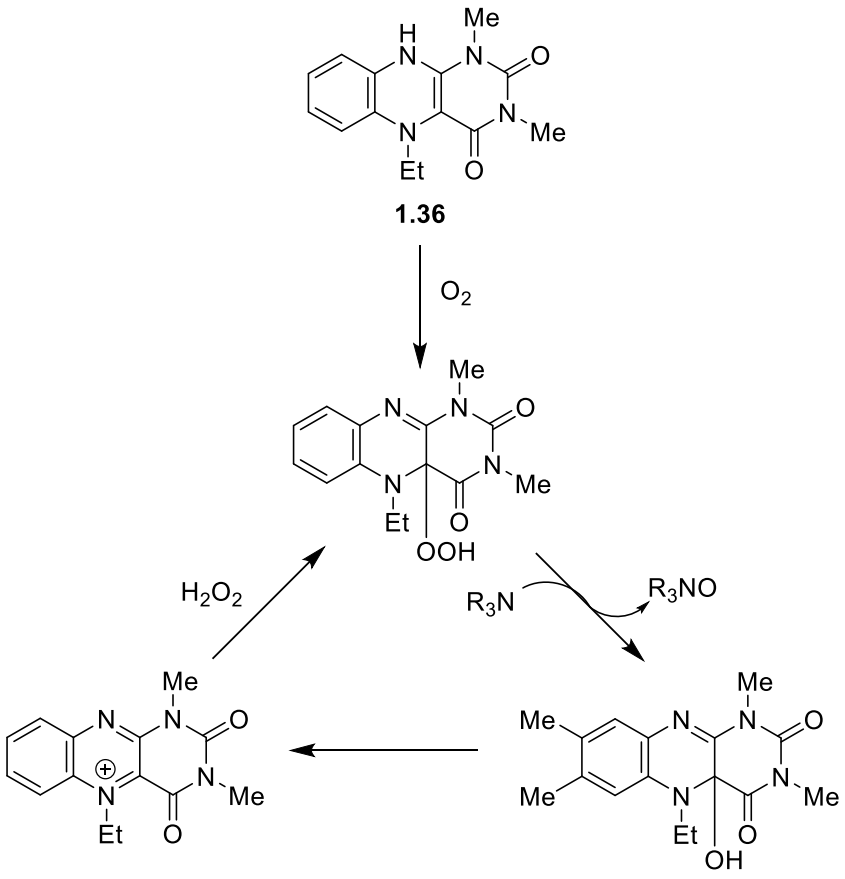
Molecular oxygen oxidations have been examined exclusively for N^5,N^{10} substituted flavinium salt catalysts. Salt **1.22** is the most commonly used catalyst for these reactions, and was used in the first oxygen oxidation of amines and sulfides catalysed by a flavinium salt by Imada.[235] A co-reductant is also required (**Scheme 1.25**), of which a variety have been found to be effective, including hydrazine,[235], zinc dust,[242] ascorbic acid,[249] and formic acid.[250]



Scheme 1.28: Preparation of dihydroalloxazine **1.36** by Bergstad *et al*.[251]

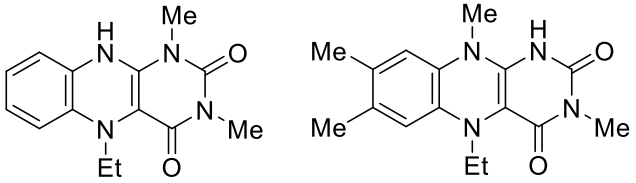
N^1,N^5 Substituted flavinium salts were first applied to catalytic oxidation reactions by Bergstad *et al*. Notably, instead of flavinium salts they elected to prepare 5,10-dihydroalloxazine **1.36** (**Scheme 1.28**).[251] Dihydroalloxazines are not catalytically active in hydrogen peroxide oxidations, but they are part of the catalytic cycle of flavins in oxygen oxidations. They are readily oxidised by air to flavin hydroperoxides. Indeed, Bergstad noted that purification of **1.36** was difficult for this very reason, and therefore air oxidation allows dihydroalloxazines to

enter the peroxide catalytic cycle (**Scheme 1.29**). The catalytic cycle is otherwise much the same as N^5,N^{10} substituted flavinium salts; involving the condensation of a diamine with alloxan to produce intermediate **1.34**, followed by methylation to afford **1.35** and reductive amination to give dihydroalloxazine **1.36**.



Scheme 1.29: Catalytic cycle for N^1,N^5 substituted 5,10-dihydroalloxazines.

Catalyst **1.36** was first prepared for use in the oxidation of tertiary amines to *N*-oxides. Rate enhancements of between 30 and 80 times were observed compared to the uncatalysed reaction.



1.36 **1.37**

Figure 1.18: Structural comparison of dihydroalloxazine **1.36** and dihydroisoalloxazine **1.37**.

Dihydroisoalloxazine **1.37** (**Figure 1.18**) displayed minimal or no activity in oxidation reactions, a curious finding as it would be expected to be oxidised by air to flavinium salt **1.22**, which has previously been shown to have high activity in oxidation reactions. This result was explained by Ménová and Cibulka, who proposed that the effect was due to differing rate-determining steps for the two reactions.[252] The rate determining step for catalysts such as dihydroalloxazine **1.36** has been shown to be elimination of water, while the rate-determining step in the reaction of dihydroisoalloxazines **1.37** is postulated to be attack of peroxide on the alloxazine cation. Basic conditions would be expected to increase the rate of reaction in the case of alloxazines, as basic conditions increase the nucleophilicity of the hydrogen peroxide. Meanwhile, acidic conditions would be expected to increase the rate of reaction in the case of isoalloxazines, as acidic conditions mean that the leaving water molecule is more easily protonated in the rate determining step. Flavinium salts generate one equivalent of acid upon first reacting with hydrogen peroxide, while dihydroalloxazines or isoalloxazines do not generate an equivalent of acid upon reaction with oxygen. Therefore, the use of a dihydroalloxazine is favoured, while the use of N^5,N^{10} substituted flavinium salts is also favoured. Furthermore, addition of acid can allow N^1,N^5 flavinium salts to be active catalysts, while addition of base allows dihydroisoalloxazines to be active catalysts. A direct comparison between dihydroalloxazines and dihydroisoalloxazines is not useful for comparing the two types of catalyst, and the same is true of a comparison between the flavinium salts. All can

display significant catalytic activity under appropriate conditions, which vary depending on the rate determining step in the catalytic cycle.

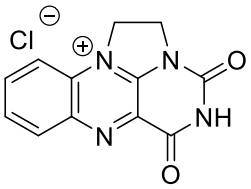
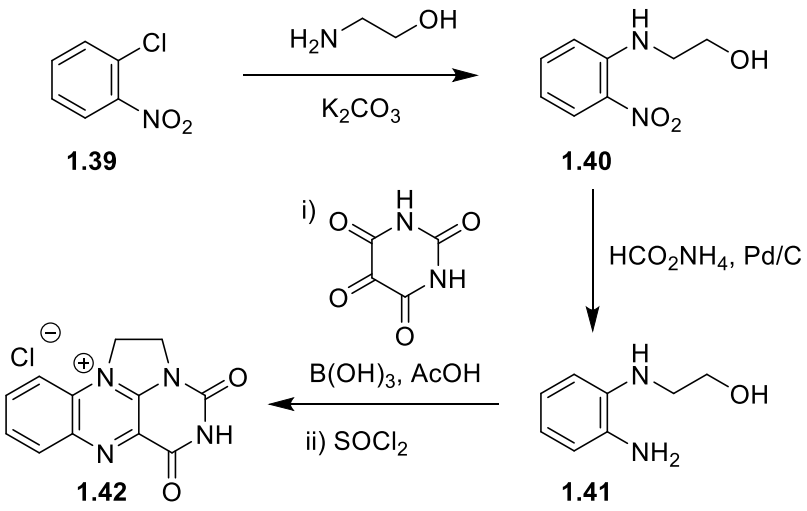


Figure 1.19: N^1,N^{10} bridged flavinium salt.

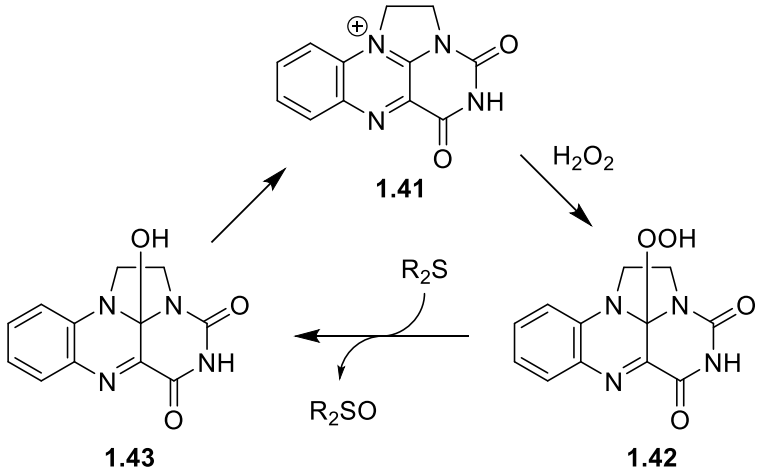
The final type of flavinium salt catalysts that have been studied are N^1,N^{10} bridged flavinium salts (**Figure 1.19**), completing the possible set of disubstitutions on N^1, N^5, and N^{10}. The first set of N^1,N^{10} bridged flavinium salts was prepared by Li *et al*. as models for natural enzyme systems.[253] Subsequently, N^1,N^{10} bridged flavinium salts as catalysts for the oxidation of sulfides were examined simultaneously by Marsh *et al*.[254] and Zurek *et al*..[255] Of the structures examined in these two papers, flavinium salt **1.42** is the simplest.



Scheme 1.30: Preparation of N^1,N^{10} bridged flavin catalysts by Marsh and Carbery.[254]

The preparation of these catalysts begins in a similar way to that of N^5,N^{10} substituted flavinium salts, with the preparation of a 1,2-diaminobenzene derivative **1.41**, and subsequent

condensation with alloxan (**Scheme 1.30**). Thionyl chloride is used to perform the cyclisation. Notably, this sequence only requires two purification steps.



Scheme 1.31: Catalytic cycle for N^1,N^{10} bridged flavinium salts.

Mechanistically, N^1,N^{10} flavinium salts behave similarly to other types of flavinium salts, but not identically. The key difference is that N^1,N^{10} bridged flavinium salts form a 10a peroxy adduct (**1.42**, **Scheme 1.31**), instead of the 4a adduct formed by N^1,N^5 and N^5,N^{10} substituted flavinium salts. This is expected, based on the fact that the positive charge in the molecule is to be found on N^{10}, not N^5.

In summary, there are a range of available flavin and flavinium salt catalysts, and a significant proportion of them have been shown to be highly active catalysts in clean, biomimetic oxidation reactions using either hydrogen peroxide or molecular oxygen as an oxidising agent. However, to date, none of these systems have been applied to the preparation of glycosyl sulfoxides.

1.15 Thesis outline

The primary aim of the research detailed within this thesis was to develop new reactions by which carbohydrate structures could be modified under aqueous conditions without the use of protecting groups using sulfur as a key reaction component.

Chapters 2, 3, and 4 detail the preparation of a variety of glycoconjugates from unprotected reducing sugars.

Chapter 2 demonstrates the use of DMC to prepare 1-cyanomethyl thioglycosides, and the indiscriminate modification of lysozyme using these reagents, allowing the production of glycoproteins.

Chapter 3 describes the preparation of glycosyl thiols in an aqueous environment through the use of DMC, and the application of the thiol-ene click reaction to selectively modify alkene-containing species with these glycosyl thiols.

Chapter 4 illustrates the use of a bifunctional alkene-alkyne linker to couple unprotected carbohydrates to thiols, including the cysteine residues of synthetic peptides.

Chapter 5 of this thesis details the preparation of glycosyl sulfoxides from carbohydrates through the use of a biomimetic flavinium salt catalyst system, and investigation of the

inhibitory activity of glycosyl sulfoxides prepared in this manner against some common glycosidases.

1.16 References

1. Ohtsubo, K.; Marth, J. D., *Cell* **2006,** *126* (5), 855-867.

2. Laine, R. A., *Glycobiology* **1994,** *4* (6), 759-767.

3. Macfarlane, M. G., *Biochem. J.* **1939,** *33* (4), 565.

4. Hooper, L. V.; Manzella, S. M.; Baenziger, J. U., *FASEB J.* **1996,** *10* (10), 1137-46.

5. Stanley, P.; Schachter, H.; Taniguchi, N., *Essentials of Glycobiology*. 2 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009.

6. Helenius, A.; Aebi, M., *Annu. Rev. Biochem.* **2004,** *73* (1), 1019-1049.

7. Breitling, J.; Aebi, M., *Cold Spring Harb. Perspect. Biol.* **2013,** *5* (8).

8. Stanley, P.; Schachter, H.; Brockhausen, I., *Essentials of Glycobiology*. 2 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009.

9. Spiro, R. G., *Glycobiology* **2002,** *12* (4), 43R-56R.

10. Pinho, S. S.; Reis, C. A., *Nat. Rev. Cancer* **2015,** *15* (9), 540-555.

11. Helenius, A., *Mol. Biol. Cell* **1994,** *5* (3), 253.

12. Leavitt, R.; Schlesinger, S.; Kornfeld, S., *J. Biol. Chem.* **1977,** *252* (24), 9018-23.

13. Gallagher, P. J.; Henneberry, J. M.; Sambrook, J. F.; Gething, M. J., *J. Virol.* **1992,** *66* (12), 7136-7145.

14. Hurtley, S. M.; Bole, D. G.; Hoover-Litty, H.; Helenius, A.; Copeland, C. S., *J. Cell. Biol.* **1989,** *108* (6), 2117-2126.

15. Miyazaki, J.-i.; Appella, E.; Zhao, H.; Forman, J.; Ozato, K., *J. Exp. Med.* **1986,** *163* (4), 856-871.

16. Roberts, P. C.; Garten, W.; Klenk, H.-D., *J. Virol.* **1993,** *67* (6), 3048-3060.

17. Ng, D. T.; Hiebert, S. W.; Lamb, R. A., *Mol. Cell. Biol.* **1990,** *10* (5), 1989-2001.

18. Landolfi, N. F.; Rich, R. R.; Cook, R. G., *J. Immunol.* **1985,** *134* (1), 423-430.

19. Struck, D. K.; Siuta, P. B.; Lane, M. D.; Lennarz, W., *J. Biol. Chem.* **1978,** *253* (15), 5332-5337.

20. Helenius, A.; Trombetta, E. S.; Hebert, D. N.; Simons, J. F., *Trends Cell Biol.* **1997,** *7* (5), 193-200.

21. Helenius, A.; Aebi, M., *Science* **2001,** *291* (5512), 2364-2369.

22. Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A., *Biochemistry* **1996,** *35* (23), 7299-7307.

23. Imperiali, B.; O'Connor, S. E., *Curr. Opin. Chem. Biol* **1999,** *3* (6), 643-649.

24. Solá, R.; Rodriguez-Martinez, J.; Griebenow, K., *Cell. Mol. Life Sci.* **2007,** *64* (16), 2133-2152.

25. van Zuylen, C. W. E. M.; Kamerling, J. P.; Vliegenthart, J. F. G., *Biochem. Biophys. Res. Commun.* **1997,** *232* (1), 117-120.

26. Arnold, U.; Ulbrich-Hofmann, R., *Biochemistry* **1997,** *36* (8), 2166-2172.

27. DeKoster, G. T.; Robertson, A. D., *Biochemistry* **1997,** *36* (8), 2323-2331.

28. Broersen, K.; Voragen, A. G. J.; Hamer, R. J.; de Jongh, H. H. J., *Biotechnol. Bioeng.* **2004,** *86* (1), 78-87.

29. Vegarud, G.; Christensen, T. B., *Biotechnol. Bioeng.* **1975,** *17* (9), 1391-1397.

30. Lawson, E. Q.; Hedlund, B.; Ericson, M. E.; Mood, D. A.; Litman, G. W.; Middaugh, R., *Arch. Biochem. Biophys.* **1983,** *220* (2), 572-575.

31. Sundaram, P. V.; Venkatesh, R., *Protein Eng. Des. Sel.* **1998,** *11* (8), 699-705.

32. Morell, A. G.; Irvine, R. A.; Sternlieb, I.; Scheinberg, I. H.; Ashwell, G., *J. Biol. Chem.* **1968,** *243* (1), 155-159.

33. Rotundo, R. F.; Rebres, R. A.; McKeown-Longo, P. J.; Blumenstock, F. A.; Saba, T. M., *Hepatology* **1998,** *28* (2), 475-485.

34. Schwartz, A. L.; Ashwell, G., *Crit. Rev. Biochem.* **1984,** *16* (3), 207-233.

35. Lowy, P. H.; Keighley, G.; Borsook, H., *Nature* **1960,** *185* (4706), 102-103.

36. Fukuda, M. N.; Sasaki, H.; Lopez, L.; Fukuda, M., *Blood* **1989,** *73* (1), 84-89.

37. Gary-Bobo, M.; Nirdé, P.; Jeanjean, A.; Morère, A.; Garcia, M., *Curr. Med. Chem.* **2007,** *14* (28), 2945-2953.

38. Braulke, T.; Bonifacino, J. S., *Biochim. Biophys. Acta Mol. Cell Res.* **2009,** *1793* (4), 605-614.

39. Zhu, Y.; Li, X.; Kyazike, J.; Zhou, Q.; Thurberg, B. L.; Raben, N.; Mattaliano, R. J.; Cheng, S. H., *J. Biol. Chem.* **2004,** *279* (48), 50336-50341.

40. Priyanka, P.; Parsons, T. B.; Miller, A.; Platt, F. M.; Fairbanks, A. J., *Angew. Chem., Int. Ed.* **2016,** *55* (16), 5058-5061.

41. Zhu, Y.; Jiang, J.-L.; Gumlaw, N. K.; Zhang, J.; Bercury, S. D.; Ziegler, R. J.; Lee, K.; Kudo, M.; Canfield, W. M.; Edmunds, T.; Jiang, C.; Mattaliano, R. J.; Cheng, S. H., *Mol. Ther.* **2009,** *17* (6), 954-963.

42. East, L.; Isacke, C. M., *Biochim. Biophys. Acta Gen. Subjects* **2002,** *1572* (2–3), 364-386.

43. Fiete, D.; Baenziger, J. U., *J. Biol. Chem.* **1997,** *272* (23), 14629-14637.

44. Largent, B. L.; Walton, K. M.; Hoppe, C. A.; Lee, Y. C.; Schnaar, R. L., *J. Biol. Chem.* **1984,** *259* (3), 1764-1769.

45. Fiete, D. J.; Beranek, M. C.; Baenziger, J. U., *Proc. Natl. Acad. Sci. U.S.A.* **1998,** *95* (5), 2089-2093.

46. Lee, S. J.; Evers, S.; Roeder, D.; Parlow, A. F.; Risteli, J.; Risteli, L.; Lee, Y. C.; Feizi, T.; Langen, H.; Nussenzweig, M. C., *Science* **2002,** *295* (5561), 1898-1901.

47. Shepherd, V. L.; Hoidal, J. R., *Am. J. Respir. Cell. Mol. Biol.* **1990,** *2* (4), 335-340.

48. Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A., *Science* **2001,** *291* (5512), 2370-2376.

49. Cresswell, P.; Bangia, N.; Dick, T.; Diedrich, G., *Immunol. Rev.* **1999,** *172* (1), 21-28.

50. Grakoui, A.; Bromley, S. K.; Sumen, C.; Davis, M. M.; Shaw, A. S.; Allen, P. M.; Dustin, M. L., *Science* **1999,** *285* (5425), 221-227.

51. Deck, M. B.; Sjölin, P.; Unanue, E. R.; Kihlberg, J., *J. Immunol.* **1999,** *162* (8), 4740-4744.

52. Pennock, N. D.; White, J. T.; Cross, E. W.; Cheney, E. E.; Tamburini, B. A.; Kedl, R. M., *Adv. Physiol. Educ.* **2013,** *37* (4), 273-283.

53. Ezekowitz, R. A.; Sastry, K.; Bailly, P.; Warner, A., *J. Exp. Med.* **1990,** *172* (6), 1785-1794.

54. Adamo, R.; Nilo, A.; Castagner, B.; Boutureira, O.; Berti, F.; Bernardes, G. J. L., *Chem. Sci.* **2013,** *4* (8), 2995-3008.

55. Morelli, L.; Poletti, L.; Lay, L., *Eur. J. Org. Chem.* **2011,** *2011* (29), 5723-5777.

56. Schneerson, R.; Barrera, O.; Sutton, A.; Robbins, J. B., *J. Exp. Med.* **1980,** *152* (2), 361-376.

57. Chu, C.; Schneerson, R.; Robbins, J. B.; Rastogi, S. C., *Infect. Immun.* **1983,** *40* (1), 245-256.

58. Gupta, R. K.; Szu, S. C.; Finkelstein, R. A.; Robbins, J. B., *Infect. Immun.* **1992,** *60* (8), 3201-3208.

59. Polotsky, V. Y.; Robbins, J. B.; Bryla, D.; Schneerson, R., *Infect. Immun.* **1994,** *62* (1), 210-214.

60. Poolman, J. T., *Expert Rev. Vaccines* **2004,** *3* (5), 597-604.

61. Johnson, M. A.; Bundle, D. R., *Chem. Soc. Rev.* **2013,** *42* (10), 4327-4344.

62. Vigerust, D. J.; Shepherd, V. L., *Trends Microbiol.* **2007,** *15* (5), 211-218.

63. Korber, B.; Gaschen, B.; Yusim, K.; Thakallapally, R.; Kesmir, C.; Detours, V., *Br. Med. Bull.* **2001,** *58* (1), 19-42.

64. Ni, J.; Song, H.; Wang, Y.; Stamatos, N. M.; Wang, L.-X., *Bioconjugate Chem.* **2006,** *17* (2), 493-500.

65. Wang, J.; Li, H.; Zou, G.; Wang, L.-X., *Org. Biomol. Chem.* **2007,** *5* (10), 1529-1540.

66. Li, H.; Wang, L.-X., *Org. Biomol. Chem.* **2004,** *2* (4), 483-488.

67. Wang, S.-K.; Liang, P.-H.; Astronomo, R. D.; Hsu, T.-L.; Hsieh, S.-L.; Burton, D. R.; Wong, C.-H., *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (10), 3690-3695.

68. Kabanova, A.; Adamo, R.; Proietti, D.; Berti, F.; Tontini, M.; Rappuoli, R.; Costantino, P., *Glycoconjugate J.* **2010,** *27* (5), 501-513.

69. Wang, L.-X., *Curr. Opin. Chem. Biol* **2013,** *17* (6), 997-1005.

70. Astronomo, R. D.; Kaltgrad, E.; Udit, A. K.; Wang, S.-K.; Doores, K. J.; Huang, C.-Y.; Pantophlet, R.; Paulson, J. C.; Wong, C.-H.; Finn, M. G.; Burton, D. R., *Chem. Biol.* **2010,** *17* (4), 357-370.

71. Doores, K. J.; Fulton, Z.; Hong, V.; Patel, M. K.; Scanlan, C. N.; Wormald, M. R.; Finn, M. G.; Burton, D. R.; Wilson, I. A.; Davis, B. G., *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (40), 17107-17112.

72. Christiansen, M. N.; Chik, J.; Lee, L.; Anugraham, M.; Abrahams, J. L.; Packer, N. H., *Proteomics* **2014,** *14* (4-5), 525-546.

73. Ragupathi, G.; Park, T. K.; Zhang, S.; Kim, I. J.; Graber, L.; Adluri, S.; Lloyd, K. O.; Danishefsky, S. J.; Livingston, P. O., *Angew. Chem., Int. Ed.* **1997,** *36* (1-2), 125-128.

74. Sabbatini, P. J.; Kudryashov, V.; Ragupathi, G.; Danishefsky, S. J.; Livingston, P. O.; Bornmann, W.; Spassova, M.; Zatorski, A.; Spriggs, D.; Aghajanian, C., *Int. J. Cancer* **2000,** *87* (1), 79-85.

75. Holmberg, L. A.; Sandmaier, B. M., *Expert Rev. Vaccines* **2004,** *3* (6), 655-663.

76. Wang, Q.; Ekanayaka, S. A.; Wu, J.; Zhang, J.; Guo, Z., *Bioconjugate Chem.* **2008,** *19* (10), 2060-2067.

77. Guo, Z.; Wang, Q., *Curr. Opin. Chem. Biol* **2009,** *13* (5–6), 608-617.

78. Toyokuni, T.; Singhal, A. K., *Chem. Soc. Rev.* **1995,** *24* (4), 231-242.

79. Franco, A., *Anticancer Agents Med. Chem.* **2008,** *8* (1), 86-91.

80. Taylor-Papadimitriou, J.; Burchell, J.; Miles, D. W.; Dalziel, M., *Biochim. Biophys. Acta Mol. Bas. Dis.* **1999,** *1455* (2–3), 301-313.

81. von Mensdorff-Pouilly, S.; Petrakou, E.; Kenemans, P.; van Uffelen, K.; Verstraeten, A. A.; Snijdewint, F. G. M.; van Kamp, G. J.; Schol, D. J.; Reis, C. A.; Price, M. R.; Livingston, P. O.; Hilgers, J., *Int. J. Cancer* **2000,** *86* (5), 702-712.

82. Zhu, J.; Wan, Q.; Ragupathi, G.; George, C. M.; Livingston, P. O.; Danishefsky, S. J., *J. Am. Chem. Soc.* **2009,** *131* (11), 4151-4158.

83. Sørensen, A. L.; Reis, C. A.; Tarp, M. A.; Mandel, U.; Ramachandran, K.; Sankaranarayanan, V.; Schwientek, T.; Graham, R.; Taylor-Papadimitriou, J.; Hollingsworth, M. A.; Burchell, J.; Clausen, H., *Glycobiology* **2006,** *16* (2), 96-107.

84. Ninkovic, T.; Kinarsky, L.; Engelmann, K.; Pisarev, V.; Sherman, S.; Finn, O. J.; Hanisch, F.-G., *Mol. Immunol.* **2009,** *47* (1), 131-140.

85. Rich, J. R.; Withers, S. G., *Nat. Chem. Biol.* **2009,** *5* (4), 206-215.

86. Villalonga, M. L.; Díez, P.; Sánchez, A.; Gamella, M.; Pingarrón, J. M.; Villalonga, R., *Chem. Rev.* **2014,** *114* (9), 4868-4917.

87. Koniev, O.; Wagner, A., *Chem. Soc. Rev.* **2015,** *44* (15), 5495-5551.

88. Villadsen, K.; Martos-Maldonado, M. C.; Jensen, K. J.; Thygesen, M. B., *ChemBioChem* **2017,** *18* (7), 574-612.

89. Goebel, W. F.; Avery, O. T., *J. Exp. Med.* **1929,** *50* (4), 521-531.

90. McBroom, C. R.; Samanen, C. H.; Goldstein, I. J., [16] Carbohydrate antigens: Coupling of carbohydrates to proteins by diazonium and phenylisothiocyanate reactions. In *Methods Enzymol.*, Academic Press: 1972; Vol. Volume 28, pp 212-219.

91. Marshall, J. J.; Rabinowitz, M. L., *Arch. Biochem. Biophys.* **1975,** *167* (2), 777-779.

92. Lasch, J.; Bessmertnaya, L.; Kozlov, L. V.; Antonov, V. K., *Eur. J. Biochem.* **1976,** *63* (2), 591-598.

93. Rosemeyer, H.; Körnig, E.; Seela, F., *Eur. J. Biochem.* **1982,** *122* (2), 375-380.

94. Gray, G. R., *Arch. Biochem. Biophys.* **1974,** *163* (1), 426-428.

95. Marsh, J. W.; Denis, J.; Wriston, J. C., *J. Biol. Chem.* **1977,** *252* (21), 7678-7684.

96. Šebela, M.; Štosová, T. á.; Havliš, J.; Wielsch, N.; Thomas, H.; Zdráhal, Z.; Shevchenko, A., *Proteomics* **2006,** *6* (10), 2959-2963.

97. Gildersleeve, J. C.; Oyelaran, O.; Simpson, J. T.; Allred, B., *Bioconjugate Chem.* **2008,** *19* (7), 1485-1490.

98. Lenders, J. P.; Crichton, R. R., *Biotechnol. Bioeng.* **1984,** *26* (11), 1343-1351.

99. Zile, J. V.; Henderson, L. A.; Baynes, J. W.; Thorpe, S. R., *J. Biol. Chem.* **1978,** *254* (9), 3547.

100. Wrodnigg, T. M.; Kartusch, C.; Illaszewicz, C., *Carbohydr. Res.* **2008,** *343* (12), 2057-2066.

101. Hojnik, C.; Müller, A.; Gloe, T.-E.; Lindhorst, T. K.; Wrodnigg, T. M., *Eur. J. Org. Chem.* **2016,** *2016* (25), 4328-4337.

102. Sakurai, K.; Miyazaki, K.; Kodera, Y.; Nishimura, H.; Shingu, M.; Inada, Y., *Glycoconjugate J.* **1997,** *14* (6), 723-728.

103. Chiumiento, A.; Dominguez, A.; Lamponi, S.; Villalonga, R.; Barbucci, R., *J. Mater. Sci. Mater. Med.* **2006,** *17* (5), 427-435.

104. Wriston, J. C., *FEBS Lett.* **1973,** *33* (1), 93-96.

105. Moczar, E., *Experientia* **1973,** *29* (12), 1576-1577.

106. Baek, W.-O.; Vijayalakshmi, M. A., *Biochim. Biophys. Acta Gen. Subjects* **1997,** *1336* (3), 394-402.

107. Solá, R. J.; Al-Azzam, W.; Griebenow, K., *Biotechnol. Bioeng.* **2006,** *94* (6), 1072-1079.

108. Solá, R. J.; Griebenow, K., *FEBS J.* **2006,** *273* (23), 5303-5319.

109. Lee, Y. C.; Stowell, C. P.; Krantz, M. J., *Biochemistry* **1976,** *15* (18), 3956-3963.

110. Stowell, C. P.; Lee, Y. C., *Biochemistry* **1980,** *19* (21), 4899-4904.

111. Pearce, O. M. T.; Fisher, K. D.; Humphries, J.; Seymour, L. W.; Smith, A.; Davis, B. G., *Angew. Chem., Int. Ed.* **2005,** *44* (7), 1057-1061.

112. Akamatsu, K.; Yamasaki, Y.; Nishikawa, M.; Takakura, Y.; Hashida, M., *J. Pharmacol. Exp. Ther.* **1999,** *290* (3), 1242-1249.

113. van Kasteren, S. I.; Campbell, S. J.; Serres, S.; Anthony, D. C.; Sibson, N. R.; Davis, B. G., *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (1), 18-23.

114. Fairbanks, A. J. The ENGases: versatile biocatalysts for the production of homogeneous N-linked glycopeptides and glycoproteins. *Chem. Soc. Rev.* [Online early access]. DOI: 10.1039/C6CS00897F. Published online: Jul 6, 2017. http://dx.doi.org/10.1039/C6CS00897F (accessed Jul 11, 2017).

115. Robbins, P. W.; Trimble, R. B.; Wirth, D. F.; Hering, C.; Maley, F.; Maley, G. F.; Das, R.; Gibson, B. W.; Royal, N.; Biemann, K., *J. Biol. Chem.* **1984,** *259* (12), 7577-83.

116. Takegawa, K.; Tabuchi, M.; Yamaguchi, S.; Kondo , A.; Kato , I.; Iwahara, S., *J. Biol. Chem.* **1995,** *270* (7), 3094-3099.

117. Kowalczyk, R.; Brimble, M. A.; Tomabechi, Y.; Fairbanks, A. J.; Fletcher, M.; Hay, D. L., *Org. Biomol. Chem.* **2014,** *12* (41), 8142-8151.

118. Wang, L.-X.; Ni, J.; Singh, S.; Li, H., *Chem. Biol.* **2004,** *11* (1), 127-134.

119. Umekawa, M.; Huang, W.; Li, B.; Fujita, K.; Ashida, H.; Wang, L.-X.; Yamamoto, K., *J. Biol. Chem.* **2008,** *283* (8), 4469-4479.

120. Fujita, M.; Shoda, S.-i.; Haneda, K.; Inazu, T.; Takegawa, K.; Yamamoto, K., *Biochim. Biophys. Acta Gen. Subjects* **2001,** *1528* (1), 9-14.

121. McIntosh, J. D.; Brimble, M. A.; Brooks, A. E. S.; Dunbar, P. R.; Kowalczyk, R.; Tomabechi, Y.; Fairbanks, A. J., *Chem. Sci.* **2015,** *6* (8), 4636-4642.

122. Huang, W.; Giddens, J.; Fan, S.-Q.; Toonstra, C.; Wang, L.-X., *J. Am. Chem. Soc.* **2012,** *134* (29), 12308-12318.

123. Oh, D.-B., *BMB Rep.* **2015,** *48* (8), 438.

124. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., *Angew. Chem.* **2002,** *114* (14), 2708-2711.

125. Tornøe, C. W.; Christensen, C.; Meldal, M., *J. Org. Chem.* **2002,** *67* (9), 3057-3064.

126. Fernandez-Gonzalez, M.; Boutureira, O.; Bernardes, G. J. L.; Chalker, J. M.; Young, M. A.; Errey, J. C.; Davis, B. G., *Chem. Sci.* **2010,** *1* (6), 709-715.

127. Lee, D. J.; Mandal, K.; Harris, P. W. R.; Brimble, M. A.; Kent, S. B. H., *Org. Lett.* **2009,** *11* (22), 5270-5273.

128. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J., *Angew. Chem., Int. Ed.* **2014,** *53* (44), 1907-11911.

129. Ullrich, M.; Liang, V.; Chew, Y. L.; Banister, S.; Song, X.; Zaw, T.; Lam, H.; Berber, S.; Kassiou, M.; Nicholas, H. R.; Götz, J., *Nat. Protocols* **2014,** *9* (9), 2237-2255.

130. Calve, S.; Witten, A. J.; Ocken, A. R.; Kinzer-Ursem, T. L., *Sci. Rep.* **2016,** *6*, 32377.

131. van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G., *Nature* **2007,** *446* (7139), 1105-1109.

132. Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G., *J. Am. Chem. Soc.* **2008,** *130* (15), 5052-5053.

133. Jongkees, S. A. K.; Umemoto, S.; Suga, H., *Chem. Sci.* **2017,** *8* (2), 1474-1481.

134. Aydillo, C.; Compañón, I.; Avenoza, A.; Busto, J. H.; Corzana, F.; Peregrina, J. M.; Zurbano, M. M., *J. Am. Chem. Soc.* **2014,** *136* (2), 789-800.

135. Davis, N. J.; Flitsch, S. L., *Tetrahedron Lett.* **1991,** *32* (46), 6793-6796.

136. Swanwick, R. S.; Daines, A. M.; Tey, L.-H.; Flitsch, S. L.; Allemann, R. K., *ChemBioChem* **2005,** *6* (8), 1338-1340.

137. Swanwick, R. S.; Daines, A. M.; Flitsch, S. L.; Allemann, R. K., *Org. Biomol. Chem.* **2005,** *3* (4), 572-574.

138. Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G., *Org. Biomol. Chem.* **2003,** *1* (21), 3642-3644.

139. Davis, B. G.; Lloyd, R. C.; Jones, J. B., *J. Org. Chem.* **1998,** *63*, 9614-9615.

140. Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G., *Angew. Chem., Int. Ed.* **2004,** *43* (7), 828-833.

141. Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G., *Angew. Chem., Int. Ed.* **2006,** *45* (24), 4007-4011.

142. McSweeney, L.; Dénès, F.; Scanlan, E. M., *Eur. J. Org. Chem.* **2016,** *2016* (12), 2080-2095.

143. Dondoni, A.; Marra, A., *Chem. Soc. Rev.* **2012,** *41* (2), 573-586.

144. Dénès, F.; Pichowicz, M.; Povie, G.; Renaud, P., *Chem. Rev.* **2014,** *114* (5), 2587-2693.

145. Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G., *Angew. Chem., Int. Ed.* **2009,** *48* (42), 7798-7802.

146. Dondoni, A.; Massi, A.; Nanni, P.; Roda, A., *Chem. Eur. J.* **2009,** *15* (43), 11444-11449.

147. Thamsen, J., *Acta Chem. Scand.* **1952,** *6* (2), 270-284.

148. Komiyama, M.; Bender, M. L., *Bioorg. Chem.* **1978,** *7* (2), 133-139.

149. Noguchi, M.; Tanaka, T.; Gyakushi, H.; Kobayashi, A.; Shoda, S. I., *J. Org. Chem.* **2009,** *74* (5), 2210-2212.

150. Huang, W.; Yang, Q.; Umekawa, M.; Yamamoto, K.; Wang, L. X., *ChemBioChem* **2010,** *11* (10), 1350-1355.

151. Tanaka, T.; Nagai, H.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Commun.* **2009,** (23), 3378-3379.

152. Tanaka, T.; Matsumoto, T.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Lett.* **2009,** *38* (5), 458-459.

153. Tanaka, T.; Huang, W. C.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Tetrahedron Lett.* **2009,** *50* (18), 2154-2157.

154. Novoa, A.; Barluenga, S.; Serba, C.; Winssinger, N., *Chem. Commun.* **2013,** *49* (69), 7608-7610.

155. Serizawa, K.; Noguchi, M.; Li, G.; Shoda, S.-i., *Chem. Lett.* **2017,** *46* (7), 1024-1026.

156. Noguchi, M.; Fujieda, T.; Huang, W. C.; Ishihara, M.; Kobayashi, A.; Shoda, S. I., *Helv. Chim. Acta* **2012,** *95* (10), 1928-1936.

157. Yoshida, N.; Noguchi, M.; Tanaka, T.; Matsumoto, T.; Aida, N.; Ishihara, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Asian. J.* **2011,** *6* (7), 1876-1885.

158. Schneider, W.; Sepp, J., *Ber. Dtsch. Chem. Ges.* **1916,** *49* (2), 2054-2057.

159. Ferrier, R. J.; Furneaux, R. H., 36 - 1,2-trans-1-Thioglycosides. In *General Methods*, BeMiller, R. L. W. N., Ed. Academic Press: 1980; pp 251-253.

160. Nicolaou, K. C.; Winssinger, N.; Pastor, J.; DeRoose, F., *J. Am. Chem. Soc.* **1997,** *119* (2), 449-450.

161. Michihata, N.; Kaneko, Y.; Kasai, Y.; Tanigawa, K.; Hirokane, T.; Higasa, S.; Yamada, H., *J. Org. Chem.* **2013,** *78* (9), 4319-4328.

162. Weng, S.-S.; Lin, Y.-D.; Chen, C.-T., *Org. Lett.* **2006,** *8* (24), 5633-5636.

163. Fügedi, P.; Garegg, P.; Lönn, H.; Norberg, T., *Glycoconjugate J.* **1987,** *4* (2), 97-108.

164. Lönn, H., *J. Carbohyd. Chem.* **1987,** *6* (2), 301-306.

165. Konradsson, P.; Udodong, U. E.; Fraser-Reid, B., *Tetrahedron Lett.* **1990,** *31* (30), 4313-4316.

166. Veeneman, G. H.; van Leeuwen, S. H.; van Boom, J. H., *Tetrahedron Lett.* **1990,** *31* (9), 1331-1334.

167. Tatai, J.; Fügedi, P., *Org. Lett.* **2007,** *9* (22), 4647-4650.

168. Kihlberg, J. O.; Leigh, D. A.; Bundle, D. R., *J. Org. Chem.* **1990,** *55* (9), 2860-2863.

169. Schmidt, R. R.; Hoffmann, M., *Tetrahedron Lett.* **1982,** *23* (4), 409-412.

170. Kahne, D.; Walker, S.; Cheng, Y.; Van Engen, D., *J. Am. Chem. Soc.* **1989,** *111* (17), 6881-6882.

171. Gargiulo, D.; Blizzard, T. A.; Nakanishi, K., *Tetrahedron* **1989,** *45* (17), 5423-5432.

172. Crich, D.; Dudkin, V., *J. Am. Chem. Soc.* **2002,** *124* (10), 2263-2266.

173. Crich, D.; Li, H.; Yao, Q.; Wink, D. J.; Sommer, R. D.; Rheingold, A. L., *J. Am. Chem. Soc.* **2001,** *123* (24), 5826-5828.

174. Crich, D.; Sun, S., *J. Org. Chem.* **1997,** *62* (5), 1198-1199.

175. Crich, D.; Wu, B.; Jayalath, P., *J. Org. Chem.* **2007,** *72* (18), 6806-6815.

176. Crich, D.; Sun, S., *J. Am. Chem. Soc.* **1997,** *119* (46), 11217-11223.

177. Crich, D.; Dudkin, V., *Org. Lett.* **2000,** *2* (24), 3941-3943.

178. Schmidt, R. R.; Stumpp, M., *Liebigs Ann. Chem.* **1984,** *1984* (4), 680-691.

179. Raghavan, S.; Kahne, D., *J. Am. Chem. Soc.* **1993,** *115* (4), 1580-1581.

180. Sliedregt, L. A. J. M.; van der Marel, G. A.; van Boom, J. H., *Tetrahedron Lett.* **1994,** *35* (23), 4015-4018.

181. Alonso, I.; Khiar, N.; Martín-Lomas, M., *Tetrahedron Lett.* **1996,** *37* (9), 1477-1480.

182. Jaramillo, C.; Corrales, G.; Fernández-Mayoralas, A., *Tetrahedron Lett.* **1998,** *39* (42), 7783-7786.

183. Gómez, A. M.; Casillas, M.; Barrio, A.; Gawel, A.; López, J. C., *Eur. J. Org. Chem.* **2008,** *2008* (23), 3933-3942.

184. Liu, J.; Huang, C.-Y.; Wong, C.-H., *Tetrahedron Lett.* **2002,** *43* (18), 3447-3448.

185. Aucagne, V.; Aversa, M. C.; Barattucci, A.; Bonaccorsi, P.; Giannetto, P.; Rollin, P.; Tatibouët, A., *J. Org. Chem.* **2002,** *67* (20), 6925-6930.

186. Aversa, M. C.; Barattucci, A.; Bilardo, M. C.; Bonaccorsi, P.; Giannetto, P.; Rollin, P.; Tatibouët, A., *J. Org. Chem.* **2005,** *70* (18), 7389-7396.

187. Aversa, M. C.; Barattucci, A.; Bonaccorsi, P., *Eur. J. Org. Chem.* **2009,** *2009* (36), 6355-6359.

188. Davies, G.; Henrissat, B., *Structure* **1995,** *3* (9), 853-859.

189. Davies, G. J.; Gloster, T. M.; Henrissat, B., *Curr. Opin. Struct. Biol.* **2005,** *15* (6), 637-645.

190. Treadway, J. L.; Mendys, P.; Hoover, D. J., *Expert Opin. Investig. Drugs* **2001,** *10* (3), 439-454.

191. Gubareva, L. V.; Kaiser, L.; Hayden, F. G., *Lancet 355* (9206), 827-835.

192. Cox, T. M.; Aerts, J. M. F. G.; Andria, G.; Beck, M.; Belmatoug, N.; Bembi, B.; Chertkoff, R.; Vom Dahl, S.; Elstein, D.; Erikson, A.; Giralt, M.; Heitner, R.; Hollak, C.; Hrebicek, M.; Lewis, S.; Mehta, A.; Pastores, G. M.; Rolfs, A.; Sa Miranda, M. C.; Zimran, A., *J. Inherit. Metab. Dis.* **2003,** *26* (6), 513-526.

193. Zitzmann, N.; Mehta, A. S.; Carrouée, S.; Butters, T. D.; Platt, F. M.; McCauley, J.; Blumberg, B. S.; Dwek, R. A.; Block, T. M., *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96* (21), 11878-11882.

194. Zechel, D. L.; Withers, S. G., *Acc. Chem. Res.* **2000,** *33* (1), 11-18.

195. Brás, N. F.; Moura-Tamames, S. A.; Fernandes, P. A.; Ramos, M. J., *J. Comput. Chem.* **2008,** *29* (15), 2565-2574.

196. Ardèvol, A.; Rovira, C., *J. Am. Chem. Soc.* **2015,** *137* (24), 7528-7547.

197. Davies, G. J.; Mackenzie, L.; Varrot, A.; Dauter, M.; Brzozowski, A. M.; Schülein, M.; Withers, S. G., *Biochemistry* **1998,** *37* (34), 11707-11713.

198. Pauling, L., *Chem. Eng. News* **1946,** *24* (10), 1375-1377.

199. Inouye, S.; Tsuruoka, T.; Ito, T.; Niida, T., *Tetrahedron* **1968,** *24* (5), 2125-2144.

200. Murao, S.; Miyata, S., *Agric. Biol. Chem.* **1980,** *44* (1), 219-221.

201. Saul, R.; Chambers, J. P.; Molyneux, R. J.; Elbein, A. D., *Arch. Biochem. Biophys.* **1983,** *221* (2), 593-597.

202. Jespersen, T. M.; Dong, W.; Sierks, M. R.; Skrydstrup, T.; Lundt, I.; Bols, M., *Angew. Chem., Int. Ed.* **1994,** *33* (17), 1778-1779.

203. Le Merrer, Y.; Fuzier, M.; Dosbaa, I.; Foglietti, M.-J.; Depezay, J.-C., *Tetrahedron* **1997,** *53* (49), 16731-16746.

204. Yoshikawa, M.; Murakami, T.; Shimada, H.; Matsuda, H.; Yamahara, J.; Tanabe, G.; Muraoka, O., *Tetrahedron Lett.* **1997,** *38* (48), 8367-8370.

205. Krause, H. P.; Keup, U.; Puls, W., *Digestion* **1982,** *23* (4), 232-238.

206. Nakamura, S.; Takahira, K.; Tanabe, G.; Morikawa, T.; Sakano, M.; Ninomiya, K.; Yoshikawa, M.; Muraoka, O.; Nakanishi, I., *Bioorg. Med. Chem. Lett.* **2010,** *20* (15), 4420-4423.

207. Ghavami, A.; Johnston, B. D.; Pinto, B. M., *J. Org. Chem.* **2001,** *66* (7), 2312-2317.

208. Gloster, T. M.; Davies, G. J., *Org. Biomol. Chem.* **2010,** *8* (2), 305-320.

209. Aoyagi, T.; Yamamoto, T.; Kojiri, K.; Morishima, H.; Nagai, M.; Hamada, M.; Takeuchi, T.; Umezawa, H., *J. Antibiot.* **1989,** *42* (6), 883-889.

210. Li, B.; Kawatkar, S. P.; George, S.; Strachan, H.; Woods, R. J.; Siriwardena, A.; Moremen, K. W.; Boons, G.-J., *ChemBioChem* **2004,** *5* (9), 1220-1227.

211. Robina, I.; Gomez-Bujedo, S.; Fernandez-Bolanos, J.; Del Pozo, L.; Demange, R.; Picasso, S.; Vogel, P., *Carbohydr. Lett.* **2000,** *3*, 389-396.

212. Poláková, M.; Šesták, S.; Lattová, E.; Petruš, L.; Mucha, J.; Tvaroška, I.; Kóňa, J., *Eur. J. Med. Chem.* **2011,** *46* (3), 944-952.

213. Colomer, J. P.; Mayordomo, M. Á. C.; De Toro, B. F.; Jiménez-Barbero, J.; Varela, O., *Eur. J. Org. Chem.* **2015,** *2015* (7), 1448-1455.

214. Lillelund, V. H.; Jensen, H. H.; Liang, X.; Bols, M., *Chem. Rev.* **2002,** *102* (2), 515-554.

215. Kakarla, R.; Dulina, R. G.; Hatzenbuhler, N. T.; Hui, Y. W.; Sofia, M. J., *J. Org. Chem.* **1996,** *61* (23), 8347-8349.

216. Breton, G. W.; Fields, J. D.; Kropp, P. J., *Tetrahedron Lett.* **1995,** *36* (22), 3825-3828.

217. Agnihotri, G.; Misra, A. K., *Tetrahedron Lett.* **2005,** *46* (47), 8113-8116.

218. Chen, M. Y.; Patkar, L. N.; Chen, H. T.; Lin, C. C., *Carbohydr. Res.* **2003,** *338* (12), 1327-1332.

219. Vincent, S. P.; Burkart, M. D.; Tsai, C. Y.; Zhang, Z.; Wong, C. H., *J. Org. Chem.* **1999,** *64* (14), 5264-5279.

220. Tsukamoto, H.; Kondo, Y., *Tetrahedron Lett.* **2003,** *44* (28), 5247-5249.

221. Chen, M. Y.; Patkar, L. N.; Lin, C. C., *J. Org. Chem.* **2004,** *69* (8), 2884-2887.

222. Huang, J.-Y.; Li, S.-J.; Wang, Y.-G., *J. Carbohyd. Chem.* **2010,** *29* (3), 142-153.

223. Wang, Q.; Wei, X.; Liao, K.; Li, H.; Meng, X.; Li, Z., *Tetrahedron Lett.* **2016,** *57* (21), 2277-2279.

224. Sato, K.; Hyodo, M.; Aoki, M.; Zheng, X.-Q.; Noyori, R., *Tetrahedron* **2001,** *57* (13), 2469-2476.

225. Legros, J.; Bolm, C., *Chem. Eur. J.* **2005,** *11* (4), 1086-1092.

226. Pitchen, P.; Dunach, E.; Deshmukh, M. N.; Kagan, H. B., *J. Am. Chem. Soc.* **1984,** *106* (26), 8188-8193.

227. Chakravarthy, R. D.; Ramkumar, V.; Chand, D. K., *Green Chem.* **2014,** *16* (4), 2190-2196.

228. Ménová, P.; Kafka, F.; Dvořáková, H.; Gunnoo, S.; Šanda, M.; Cibulka, R., *Adv. Synth. Catal.* **2011,** *353* (6), 865-870.

229. Bahrami, K.; Khodaei, M. M.; Sheikh Arabi, M., *J. Org. Chem.* **2010,** *75* (18), 6208-6213.

230. Gu, X.; Li, X.; Chai, Y.; Yang, Q.; Li, P.; Yao, Y., *Green Chem.* **2013,** *15* (2), 357-361.

231. Kumar, S.; Verma, S.; Jain, S. L.; Sain, B., *Tetrahedron Lett.* **2011,** *52* (26), 3393-3396.

232. Murahashi, S. I.; Oda, T.; Masui, Y., *J. Am. Chem. Soc.* **1989,** *111* (13), 5002-5003.

233. Dijkman, W. P.; de Gonzalo, G.; Mattevi, A.; Fraaije, M. W., *Appl. Microbiol. Biotechnol.* **2013,** *97* (12), 5177-5188.

234. Schrecker, A. W.; Kornberg, A., *J. Biol. Chem.* **1950,** *182* (2), 795-803.

235. Imada, Y.; Iida, H.; Ono, S.; Murahashi, S. I., *J. Am. Chem. Soc.* **2003,** *125* (10), 2868-2869.

236. Ball, S.; Bruice, T. C., *J. Am. Chem. Soc.* **1980,** *102* (21), 6498-6503.

237. Krueger, S. K.; Williams, D. E., *Pharmacol. Ther.* **2005,** *106* (3), 357-387.

238. Ghisla, S.; Massey, V., *Eur. J. Biochem.* **1989,** *181* (1), 1-17.

239. Mager, H. I. X.; Tu, S.-C., *Tetrahedron* **1994,** *50* (23), 6759-6766.

240. Kotoucova, H.; Strnadova, I.; Kovandova, M.; Chudoba, J.; Dvorakova, H.; Cibulka, R., *Org. Biomol. Chem.* **2014,** *12* (13), 2137-2142.

241. Mazzini, C.; Lebreton, J.; Furstoss, R., *J. Org. Chem.* **1996,** *61* (1), 8-9.

242. Imada, Y.; Iida, H.; Murahashi, S.-I.; Naota, T., *Angew. Chem., Int. Ed.* **2005,** *44* (11), 1704-1706.

243. Imada, Y.; Iida, H.; Ono, S.; Masui, Y.; Murahashi , S.-I., *Chem. Asian. J.* **2006,** *1* (1-2), 136-147.

244. Baxová, L.; Cibulka, R.; Hampl, F., *J. Mol. Catal. A: Chem.* **2007,** *277* (1–2), 53-60.

245. Minidis, A. B. E.; Bäckvall, J. E., *Chem. Eur. J.* **2001,** *7* (1), 297-302.

246. Jurok, R.; Cibulka, R.; Dvořáková, H.; Hampl, F.; Hodačová, J., *Eur. J. Org. Chem.* **2010,** *2010* (27), 5217-5224.

247. Murahashi, S.-I.; Ono, S.; Imada, Y., *Angew. Chem., Int. Ed.* **2002,** *41* (13), 2366-2368.

248. Imada, Y.; Ohno, T.; Naota, T., *Tetrahedron Lett.* **2007,** *48* (6), 937-939.

249. Imada, Y.; Kitagawa, T.; Wang, H.-K.; Komiya, N.; Naota, T., *Tetrahedron Lett.* **2013,** *54* (7), 621-624.

250. Murahashi, S.-I.; Zhang, D.; Iida, H.; Miyawaki, T.; Uenaka, M.; Murano, K.; Meguro, K., *Chem. Commun.* **2014,** *50* (71), 10295-10298.

251. Bergstad, K.; Bäckvall, J.-E., *J. Org. Chem.* **1998,** *63* (19), 6650-6655.

252. Ménová, P.; Cibulka, R., *J. Mol. Catal. A: Chem.* **2012,** *363–364*, 362-370.

253. Li, W. S.; Zhang, N.; Sayre, L. M., *Tetrahedron* **2001,** *57* (21), 4507-4522.

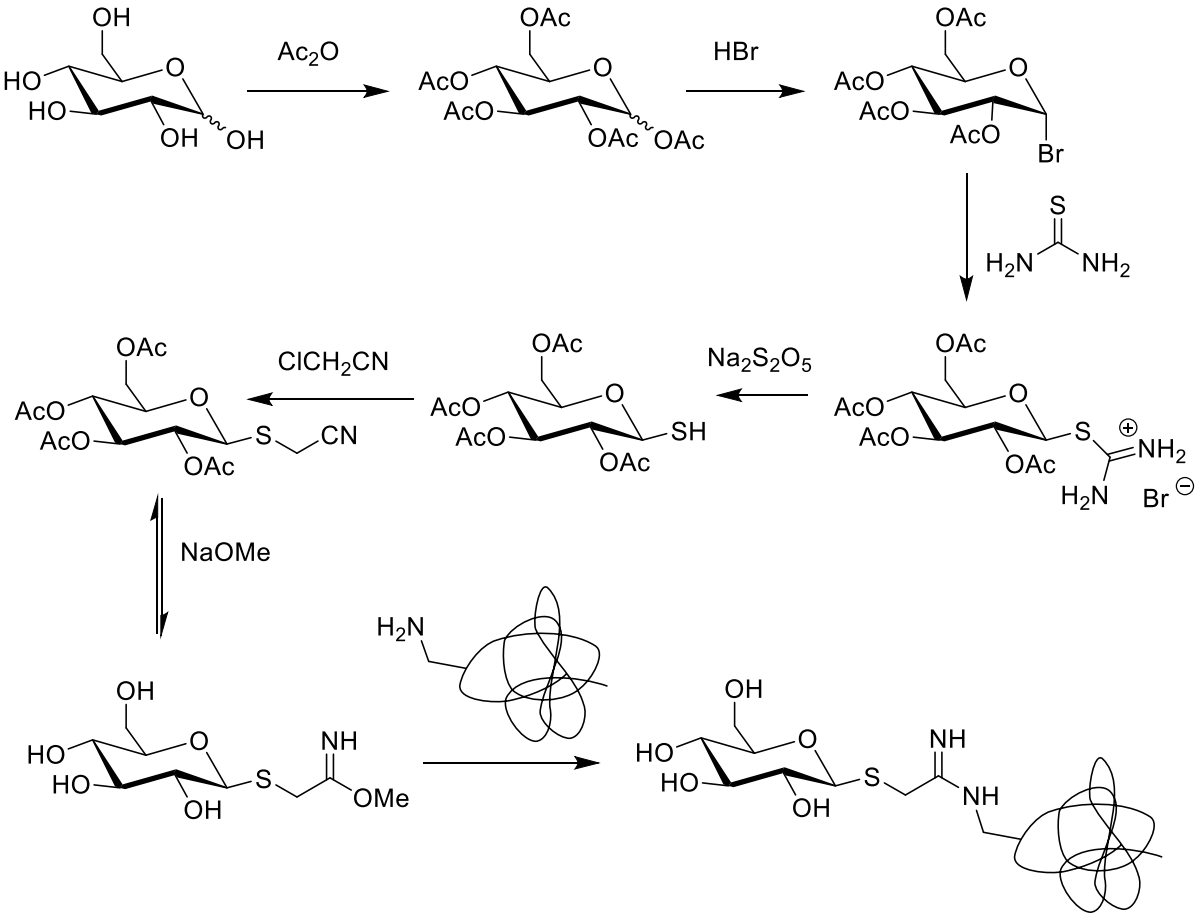
254. Marsh, B. J.; Carbery, D. R., *Tetrahedron Lett.* **2010,** *51* (17), 2362-2365.

255. Žurek, J.; Cibulka, R.; Dvořáková, H.; Svoboda, J., *Tetrahedron Lett.* **2010,** *51* (7), 1083-1086.

Chapter 2: The aqueous preparation of cyanomethyl thioglycosides and their use in the preparation of glycoproteins

2.1 Cyanomethyl thioglycosides

As discussed previously (**Section 1.7.1**), cyanomethyl thioglycosides are an effective way of indiscriminately glycosylating a protein while still maintaining the entire intact carbohydrate structure and are additionally completely selective for lysine residues. However, the preparation of cyanomethyl thioglycosides and their conjugation to protein requires a number of synthetic steps, and is therefore difficult to carry out with larger oligosaccharide substrates.[1]

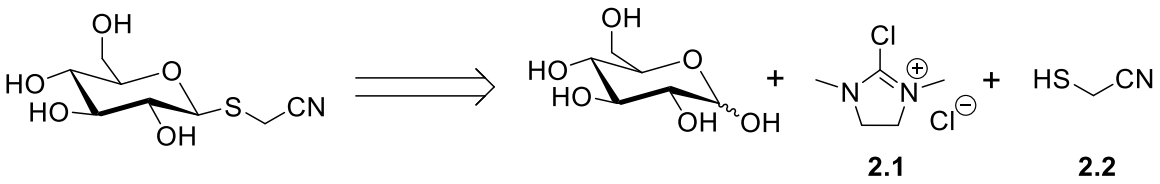


Scheme 2.1: Typical route for the preparation and conjugation of cyanomethyl thioglycosides.[1]

Currently, the preparation of cyanomethyl glycosides and their conjugation to proteins is an extended process (**Scheme 2.1**). The carbohydrate is first acetylated with acetic anhydride, then converted to the glycosyl bromide. The glycosyl bromide is reacted with thiourea to produce the isothiouronium salt, which is then cleaved with sodium metabisulfite to give the glycosyl thiol, which is subsequently converted to the cyanomethyl thioglycoside by reaction with chloroacetonitrile. Simultaneous acetate deprotection and equilibration to form the 2-imino-2-methoxyethyl 1-thioglycoside (IME thioglycoside) is followed by conjugation to a protein using the crude reaction mixture. While this method is relatively effective for smaller oligosaccharides, the multiple protection and deprotection steps become extremely challenging for larger oligosaccharides. A method which bypasses the requirement for protection and deprotection would facilitate the preparation of glycoproteins using larger oligosaccharides *via* IME thioglycoside chemistry.

2.2 The use of DMC for preparing cyanomethyl thioglycosides

As discussed previously (**Section 1.8**), DMC (**2.1**) is a highly effective reagent for the modification of the anomeric centre of unprotected carbohydrates with strongly nucleophilic reagents in an aqueous environment.[2, 3] Thiols are good nucleophiles, and indeed have been used in DMC-mediated reactions in the past.[3-5] It therefore seemed logical that the use of mercaptoacetonitrile (**2.2**) as a nucleophile together with DMC activation of the anomeric centre could serve to generate cyanomethyl thioglycosides, which could then be used as reagents for lysine modification without requiring protection of the carbohydrate at any stage (**Scheme 2.2**).

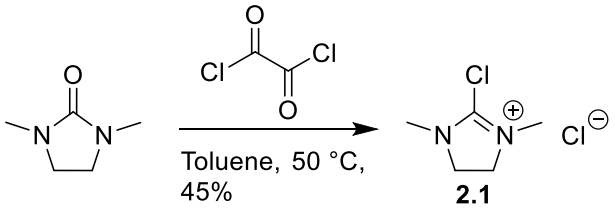


Scheme 2.2: Retrosynthesis of cyanomethyl thioglycosides using DMC.

In addition to establishing that thiols can be effective nucleophiles for reaction with sugars in a DMC-mediated process, Shoda[3] and Winssinger[5] furthermore both observed that more acidic thiols display greater reactivity in the DMC reaction, due to the tendency of more basic thiols to react with DMC preferentially over the carbohydrate anomeric hydroxyl group. Mercaptoacetonitrile would be expected to be relatively acidic due to the presence of the electron-withdrawing nitrile group, providing further support for the potential reaction of mercaptoacetonitrile with DMC-activated sugars.

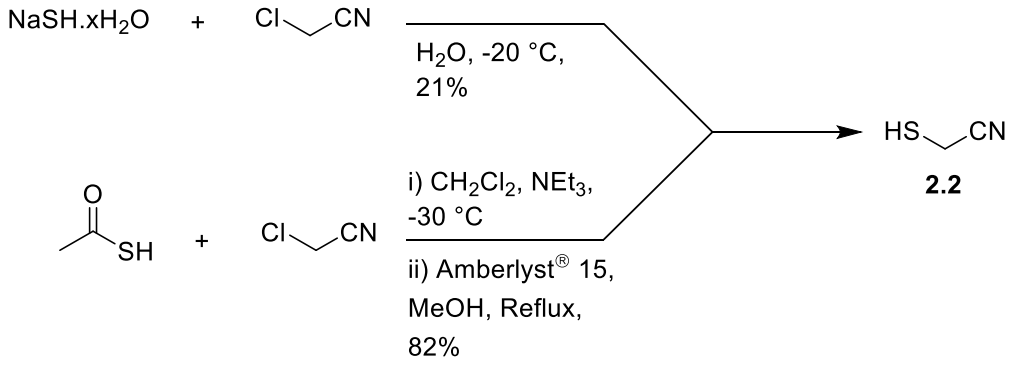
2.3 Reagent preparation: DMC and mercaptoacetonitrile

Both DMC[6] and mercaptoacetonitrile[7, 8] have been previously prepared, and so both were prepared according to the relevant literature procedures.



Scheme 2.3: Preparation of DMC.

DMC (**2.1**) was prepared in 45% yield by reaction of 1,3-dimethylimidazolidinone with oxalyl chloride at 50 °C.[6]



Scheme 2.4: Preparation of mercaptoacetonitrile.

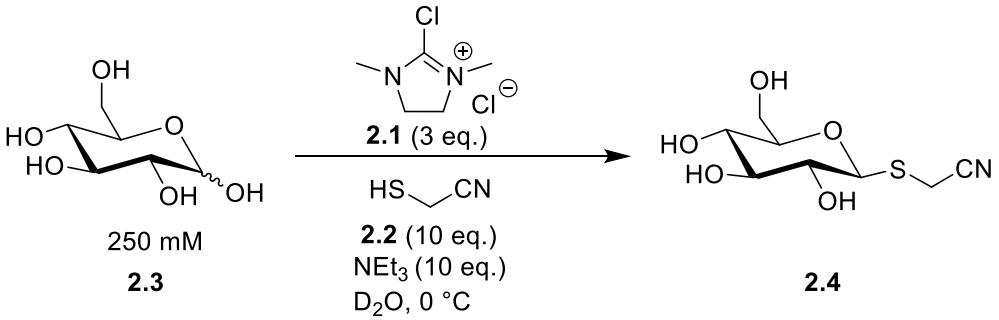
At first, mercaptoacetonitrile (**2.2**) was prepared by reaction of chloroacetonitrile with sodium hydrosulfide hydrate at -20 °C according to the method of Mathias and Shimanski (**Scheme 2.4**).[7] While this reaction does produce mercaptoacetonitrile, it is low-yielding and inconsistent, presumably due to dimerization *via* reaction of mercaptoacetonitrile product with chloroacetonitrile starting material, dimerization *via* disulfide formation, and polymerisation of mercaptoacetonitrile.[9] Careful control of the reaction temperature was required, as low temperatures resulted in the freezing of the reaction mixture, and high temperatures resulted in minimal product formation. Rapid addition of chloroacetonitrile is also required, but the reaction is exothermic, again resulting in challenges for temperature control. The removal of significant impurities was also difficult, as mercaptoacetonitrile is unstable at high temperatures, meaning that distillation is not straightforward.

An alternative route for the preparation of mercaptoacetonitrile was therefore examined. Gaumont *et al*. prepared mercaptoacetonitrile by reaction of thioacetic acid with chloroacetonitrile, using triethylamine as a base to form a thioester which was then deprotected under acidic conditions to afford mercaptoacetonitrile (**Scheme 2.4**).[8] This method has the advantage of at no point generating a reactive thiolate anion, and therefore being significantly less likely to produce byproducts. Mercaptoacetonitrile was prepared by this method in 82%

yield. Furthermore, the reaction was significantly cleaner, making purification far less challenging.

2.4 Investigation of conditions for the preparation of cyanomethyl thioglycosides

In order to investigate the conditions for the preparation of unprotected cyanomethyl thioglycosides from reducing sugars using DMC, glucose was selected as a model substrate for 2-hydroxy sugars. As a starting point, standard DMC activation conditions were used (**Scheme 2.5**).[3]



Scheme 2.5: Conditions for the DMC activation of glucose and subsequent nucleophilic substitution with mercaptoacetonitrile.

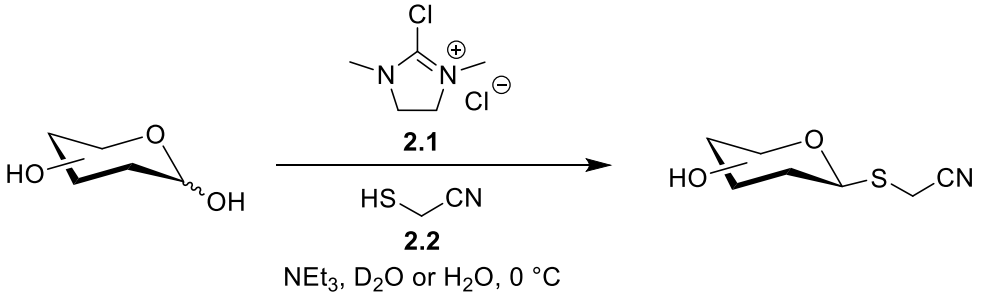
These conditions resulted in a small but noticeable conversion of 30% to cyanomethyl thioglucoside **2.4**, as determined by ^{1}H NMR spectroscopy, with the remainder of the material being glucose **2.3**. The order of addition of reagents is typically important in DMC-mediated reactions, and in this case glucose **2.3**, triethylamine, and mercaptoacetonitrile were mixed in D_2O, followed by the portionwise addition of DMC. Notably, no 1,6-anhydroglucose was produced, which is a common byproduct of reactions of 2-hydroxy sugars with DMC.[10]

Variation of the order of addition of the reagents significantly affected the course of the reaction. Mixing glucose **2.3**, DMC, and mercaptoacetonitrile in D_2O and then adding triethylamine resulted in 39% conversion to **2.4**, a slight improvement, with the remainder again being glucose **2.3**. Alternatively, addition of the reactants in the order glucose **2.3**, triethylamine, DMC, and then adding the mercaptoacetonitrile resulted in 40% conversion to **2.4**, but also produced 23% of 1,6-anhydroglucose. The use of a 1:1 mixture of D_2O:acetonitrile as the solvent resulted in no conversion to product, and only glucose **2.3** starting material was observed. However, the addition of DMC to the reaction mixture last as a solution in mercaptoacetonitrile resulted in a noticeable increase in conversion to 53%. It was further determined upon repetition of the reaction that dropwise addition of the DMC-mercaptoacetonitrile solution was critical to obtain a high conversion, with reduced conversion to product **2.4** observed when the DMC-mercaptoacetonitrile solution was added more rapidly.

From these results, it would appear that mercaptoacetonitrile is unstable under the basic aqueous conditions of the DMC-mediated reaction, and therefore adding it to the reaction as late as possible results in a greater conversion to product, as there is less time for the mercaptoacetonitrile to decompose.

In order to further improve the conversion to product, the number of equivalents of reagents was doubled, to 6 equivalents of DMC, 20 equivalents of mercaptoacetonitrile, and 20 equivalents of triethylamine. Pleasingly, this resulted in a conversion of glucose **2.3** to cyanomethyl thioglycoside **2.4** of 83%, and an isolated yield of product **2.4** of 65%. Furthermore, when water was used as a solvent a conversion of glucose **2.3** to product **2.4** of 91% and an isolated yield of 89% were obtained. The variation in yield between the deuterated

and non-deuterated solvent is most likely due to experimental variation rather than any meaningful effect given the apparently random nature of the variation. As discussed previously, rate of addition of the reagents had a strong effect on the conversion to product, and subtle changes in the rate of addition could explain variations in conversion to product



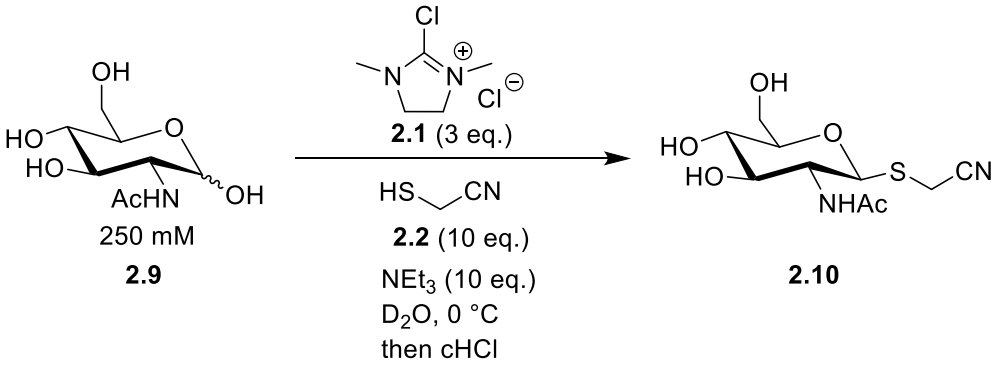
Scheme 2.6: Testing of the substrate scope of DMC-mediated cyanomethyl thioglycoside synthesis.

Substrate	Product	Solvent	Isolated Yield
Glucose	**2.4**	H_2O	89%
		D_2O	65%
Mannose	**2.5**	H_2O	78%
		D_2O	79%
Galactose	**2.6**	H_2O	41%
		D_2O	26%
Lactose	**2.7**	H_2O	65%
		D_2O	42%
Maltotriose	**2.8**	D_2O	90%

Table 2.1: Substrate scope of DMC-mediated cyanomethyl thioglycoside synthesis.

With reaction conditions for 2-hydroxy sugars in hand, attention then turned to the substrate scope of the reaction. The reaction was shown to be effective and gave the corresponding cyanomethyl thioglycosides in moderate to high yield with a variety of 2-hydroxy sugars, using both D_2O and H_2O as solvent (**Table 2.1**). When D_2O was used a solvent a degree of H/D exchange was observed at the aglycon CH_2, presumably because of the increased acidity of the

hydrogens at this position due to the adjacent nitrile group. Cyanomethyl thiogalactoside **2.6** was produced in relatively low yield, an observation consistent with previous DMC-mediated reactions of galactose using thiols as a nucleophile.[4, 5] Substituent orientation in carbohydrates can have an effect on the rate of substitution reactions at the anomeric centre.[11] This effect is proposed to be due to different electron donating abilities of axial and equatorial hydroxyl groups.[12] It therefore seems reasonable that a similar effect is in play here, as all proposed DMC reaction mechanisms proceed *via* strained bicyclic intermediates.



Scheme 2.7: Conditions for the DMC reaction of GlcNAc and mercaptoacetonitrile.

Attention then turned to the application of the DMC-mediated reaction to 2-acetamido sugars. In a standard DMC-mediated reaction for 2-acetamido sugars, the sugar is first activated by DMC in order to form a 1,2-oxazoline, which is then opened by addition of acid in the presence of the nucleophile (**Scheme 2.7**).[6] Unfortunately, these conditions gave a conversion of *N*-acetylglucosamine (GlcNAc) **2.9** to cyanomethyl thioglycoside **2.10** of only 16%, with the remainder of the material being the starting material. Further investigation revealed that a comparable conversion to **2.10** was achieved when the acid was not added, in which case the remainder of the material was found to be 1,2-oxazoline. As discussed previously, mercaptoacetonitrile is unstable under the basic aqueous conditions of the DMC reaction, and therefore one conclusion is that all of the cyanomethyl thioglycoside **2.10** that forms does so by direct attack of mercaptoacetonitrile on the DMC-sugar intermediate, rather than by opening

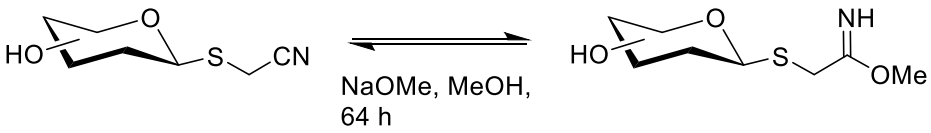
of the 1,2-oxazoline; i.e., by the time acid is added all the mercaptoacetonitrile has decomposed. Furthermore, the addition of mercaptoacetonitrile as a solution in concentrated HCl to preformed oxazoline did not result in any product, a result which can be explained by the reduced nucleophilicity of the thiol species found under acidic conditions as compared to the thiolate anion that is present under basic conditions.

Due to the inevitability of significant oxazoline formation by the rapid intramolecular attack of the 2-acetamido group at the anomeric centre, and the mutual incompatibility of oxazoline ring opening and high nucleophilicity of mercaptoacetonitrile, a viable reaction for the formation of cyanomethyl thioglycosides of 2-acetamido sugars by activation with DMC and reaction with mercaptoacetonitrile was deemed unlikely. Therefore at this point 2-acetamido sugars as substrates were abandoned to focus on the development of the more effective reaction of 2-hydroxy sugars.

2.5 Preparation of IME thioglycosides

With a selection of cyanomethyl thioglycosides of 2-hydroxy sugars in hand, attention then turned to the formation of the corresponding IME thioglycosides. IME Thioglycosides are typically prepared by the reaction of the corresponding cyanomethyl glycosides with sodium methoxide in an equilibrium process, which typically yields around 40% conversion to IME thioglycoside (**Scheme 2.8**).[1] In other examples of the preparation of IME thioglycosides the sodium methoxide also serves to deprotect the sugar, but in the case of unprotected cyanomethyl thioglycosides deprotection is not necessary. The IME thioglycoside-

cyanomethyl thioglycoside mixture is used typically as a crude mixture to modify proteins; the remaining cyanomethyl thioglycoside is not removed.



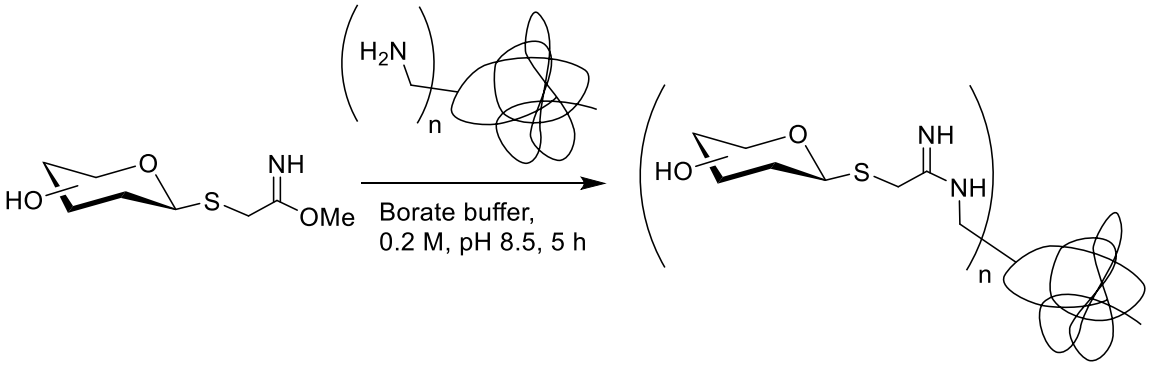
Scheme 2.8: Preparation of IME thioglycosides.

Cyanomethyl Thioglycoside	**Product**	**Conversion**
Glucoside 2.4	**2.11**	35%
Mannoside 2.5	**2.12**	22%
Galactoside 2.6	**2.13**	50%
Lactoside 2.7	**2.14**	85%
Maltotrioside 2.8	**2.15**	60%

Table 2.2: Preparation of IME thioglycosides.

As expected, in all cases the preparation of IME thioglycosides (**Scheme 2.8**) resulted in the formation of a mixture of the IME thioglycoside and the corresponding cyanomethyl thioglycoside starting material (**Table 2.2**). Some variation in conversion to IME thioglycoside was observed, a result attributed to subtly different equilibrium positions for each sugar.

2.6 Protein modification using IME thioglycosides



Scheme 2.9: Protein modification with IME thioglycosides.

With a selection of IME thioglycosides in hand, attention turned to the modification of a model protein. Modification of protein lysines with IME reagents is typically carried out at pH 8.5 in a sodium borate buffer (**Scheme 2.9**); the amidination reaction being more effective at higher pH.[13] The use of an even higher pH results in greater modification, but also risks denaturing the protein.

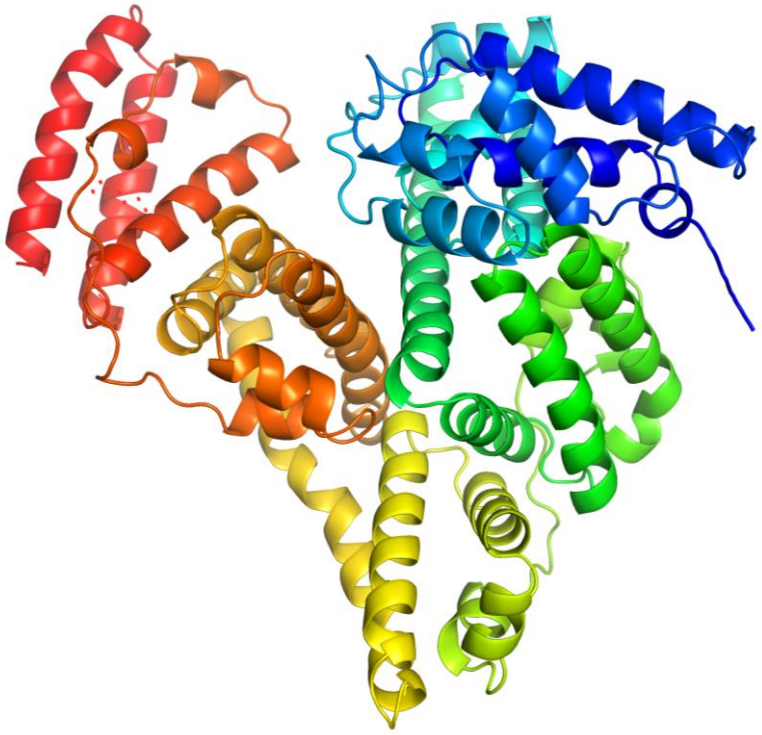


Figure 2.1: Structure of BSA.[14]

BSA (**Figure 2.1**) was the first protein examined, and unfortunately did not prove suitable. While it appeared that modification had occurred, MS analysis did not provide suitable characterisation data. Amidination produces a statistical distribution of modified protein: not every lysine residue is modified and the exact modification positions are random, leading to a wide range of possible product masses. This problem was compounded by the relatively high molecular weight of BSA (approximately 66.4 kDa), which also results in a high number of lysine residues, 59 in all, any of which can be modified. As a result, the data obtained upon MS analysis of the reaction of BSA with IME thioglucoside **2.4** did not clearly provide evidence of conjugate formation, although it appeared that all BSA had been consumed.

Figure 2.2: Structure of hen egg white lysozyme.[15]

With this methodological limitation in mind, a smaller protein was chosen for exemplification. Hen egg white lysozyme is a simple protein with a molecular weight of approximately 14.3 kDa and only 6 lysine residues, meaning that characterisation of any conjugate generated was expected to be significantly simpler.

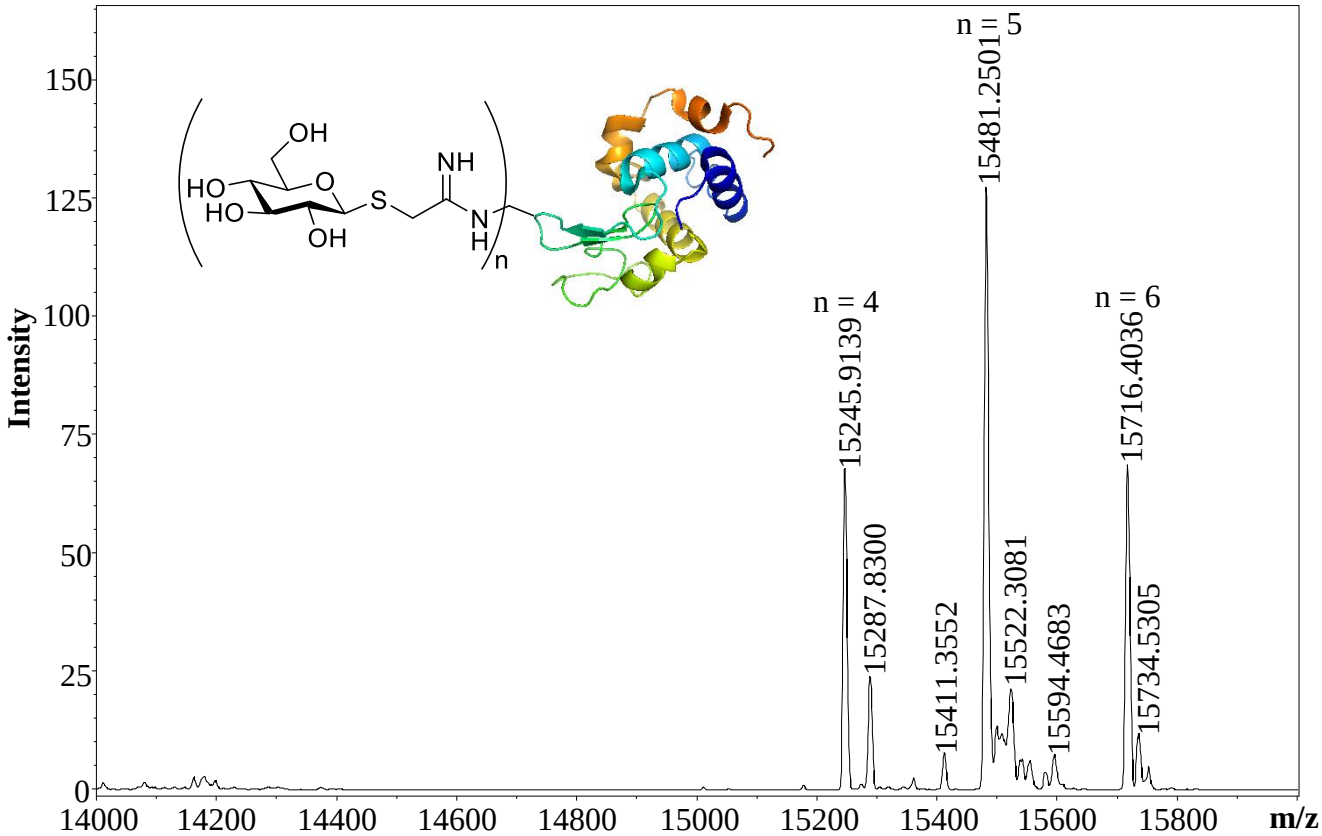
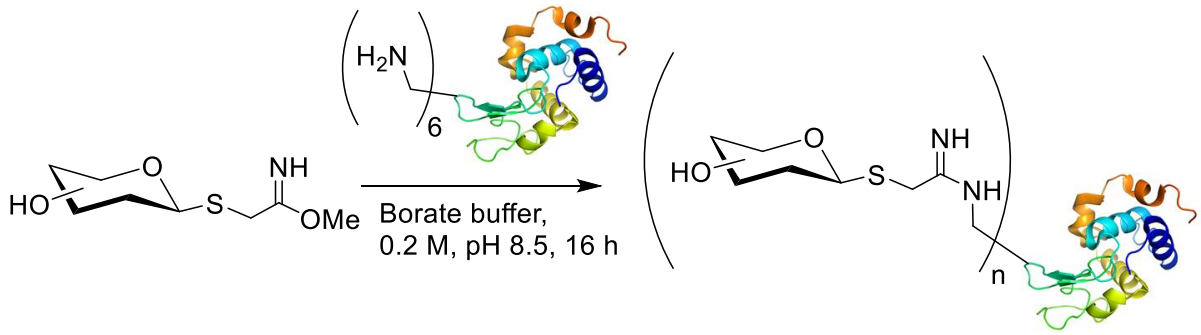


Figure 2.3: Deconvoluted HRMS of lysozyme modified with glucose reagent **2.11**.

Gratifyingly, this proved to be the case, and reaction of various IME thioglycosides with lysozyme resulted in significant modification. In all cases 50 µL of the IME thioglycoside equilibrium mixture, with a combined sugar concentration of 0.1 M, was added to 0.070 µmol of lysozyme in pH 8.5 borate buffer (**Scheme 2.10**), representing between 2 and 10 equivalents of IME glycoside per lysine residue, depending on the equilibrium of the cyanomethyl thioglycoside-IME thioglycoside equilibrium for the sugar in question. Characterisation of the modified lysozyme was carried out by MS analysis (**Figure 2.3**), allowing identification and quantification of all modified lysozyme species present.



Scheme 2.10: Modification of lysozyme with IME thioglycosides.

IME Thioglycoside	Number of Sugar Units Attached							
	0	**1**	**2**	**3**	**4**	**5**	**6**	**Mean**
Glucoside 2.11	-	-	-	-	26%	48%	26%	5.0
Mannoside 2.12	9%	33%	38%	18%	2%	-	-	1.7
Galactoside 2.13	-	-	-	1%	28%	46%	25%	5.0
Lactoside 2.14	3%	9%	27%	35%	16%	9%	-	2.8
Maltotrioside 2.15	4%	26%	43%	21%	7%	<1%	-	2.0

Table 2.3: Modification of lysozyme with IME thioglycosides.

In all cases, multiple sugar sugars units were observed to be conjugated to lysozyme (**Table 2.3**). Glucose and galactose IME thioglycosides **2.11** and **2.13** both modified the majority of the lysozyme's six lysine residues, while the use of the mannose, lactose, and maltotriose IME thioglycosides **2.12**, **2.14**, and **2.15** resulted in at least some modification in all cases. Modification with mannoside **2.12** resulted in reduced modification relative to glucoside **2.11** and galactoside **2.13**, an observation for which the most likely explanation is the reduced conversion of cyanomethyl thioglycoside to IME thioglycoside. This explanation was supported by the increased conjugation observed upon the addition of further equivalents of IME thioglycoside; a second addition of IME thioglycoside resulted in an increase in the average number of attached sugars to 3.8.

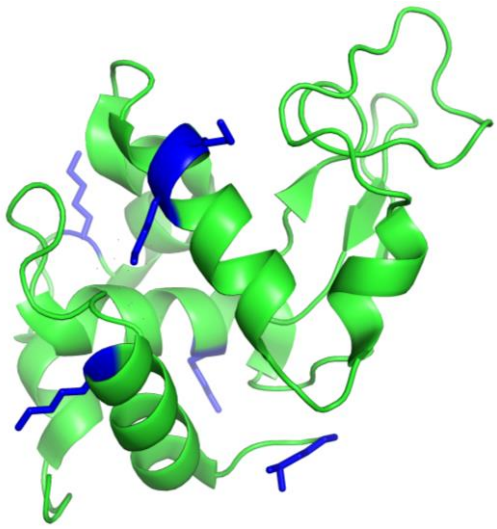


Figure 2.4: Structure of lysozyme with lysine residues highlighted in blue.

Meanwhile, the lower modification observed for lactoside **2.14** and maltotrioside **2.15** may to be due to greater steric hindrance resulting from the larger glycans. As the modification reaction competes with imidate hydrolysis, any reduction in reaction rate is accompanied by a corresponding drop in conjugation yield. This proposal is supported by the progressive decline in the observed average number of sugars attached to the protein as the sugar itself becomes larger. While lysozyme is not a large protein and the lysine residues are relatively exposed (**Figure 2.4**), as the sugar unit becomes larger it will still become progressively more difficult to attach.

2.7 Conclusions

A procedure was developed for the preparation of cyanomethyl thioglycosides from unprotected reducing sugars in water, and was applied to a range of 2-hydroxy sugars. The cyanomethyl thioglycosides produced by this reaction were then converted to the corresponding IME thioglycosides and conjugated to hen egg white lysozyme, demonstrating the utility of this method for the preparation of glycoproteins. The method developed is

significantly less complex than existing procedures, and should enable the more widespread use of IME thioglycosides for the preparation of glycoproteins.

2.8 References

1. Lee, Y. C.; Stowell, C. P.; Krantz, M. J., *Biochemistry* **1976,** *15* (18), 3956-3963.

2. Tanaka, T.; Nagai, H.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Commun.* **2009,** (23), 3378-3379.

3. Tanaka, T.; Matsumoto, T.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Lett.* **2009,** *38* (5), 458-459.

4. Yoshida, N.; Noguchi, M.; Tanaka, T.; Matsumoto, T.; Aida, N.; Ishihara, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Asian. J.* **2011,** *6* (7), 1876-1885.

5. Novoa, A.; Barluenga, S.; Serba, C.; Winssinger, N., *Chem. Commun.* **2013,** *49* (69), 7608-7610.

6. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J., *Angew. Chem., Int. Ed.* **2014,** *53* (44), 1907-11911.

7. Mathias, E.; Shimanski, M., *J. Chem. Soc., Chem. Commun.* **1981,** (11), 569-570.

8. Gaumont, A. C.; Wazneh, L.; Denis, J. M., *Tetrahedron* **1991,** *47* (27), 4927-4940.

9. Arrhenius, G.; Bladridge, K. K.; Richards-Gross, S.; Siegel, J. S., *J. Org. Chem.* **1997,** *62* (16), 5522-5525.

10. Tanaka, T.; Huang, W. C.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Tetrahedron Lett.* **2009,** *50* (18), 2154-2157.

11. Namchuk, M. N.; McCarter, J. D.; Becalski, A.; Andrews, T.; Withers, S. G., *J. Am. Chem. Soc.* **2000,** *122* (7), 1270-1277.

12. Jensen, H. H.; Bols, M., *Org. Lett.* **2003,** *5* (19), 3419-3421.

13. Stowell, C. P.; Lee, Y. C., *Biochemistry* **1980,** *19* (21), 4899-4904.

14. Bujacz, A., *Acta Crystallogr. D Struct. Biol.* **2012,** *68* (10), 1278-1289.

15. Weiss, M. S.; Palm, G. J.; Hilgenfeld, R., *Acta Crystallogr. D Struct. Biol.* **2000,** *56* (Pt 8), 952-8.

Chapter 3: The aqueous preparation of glycosyl thiols and their use in the thiol-ene click reaction

3.1 The site-selective preparation of glycoproteins and glycopeptides

As discussed in **Section 1.7.2**, the preparation of glycoproteins and glycopeptides in a site-selective manner is typically more useful than non-specific approaches, as the biological activity of these glycoconjugates can be strongly tied to the position(s) of glycosylation and the structures of attached glycans. However, the most significant drawback associated with the site-selective preparation of glycoproteins is that this is far more synthetically challenging than making indiscriminately glycosylated glycoconjugates. In particular, the chemical synthesis of glycoproteins and glycopeptides almost invariably requires protection and deprotection of the sugar, a process that is particularly complex when applied to larger oligosaccharides. In recent years some attempts have been made to overcome this particular hurdle, and one of the most promising of these is the use of DMC to modify unprotected carbohydrates (**Section 1.8**).

3.2 The search for alternative nucleophiles for reaction with DMC-activated sugars

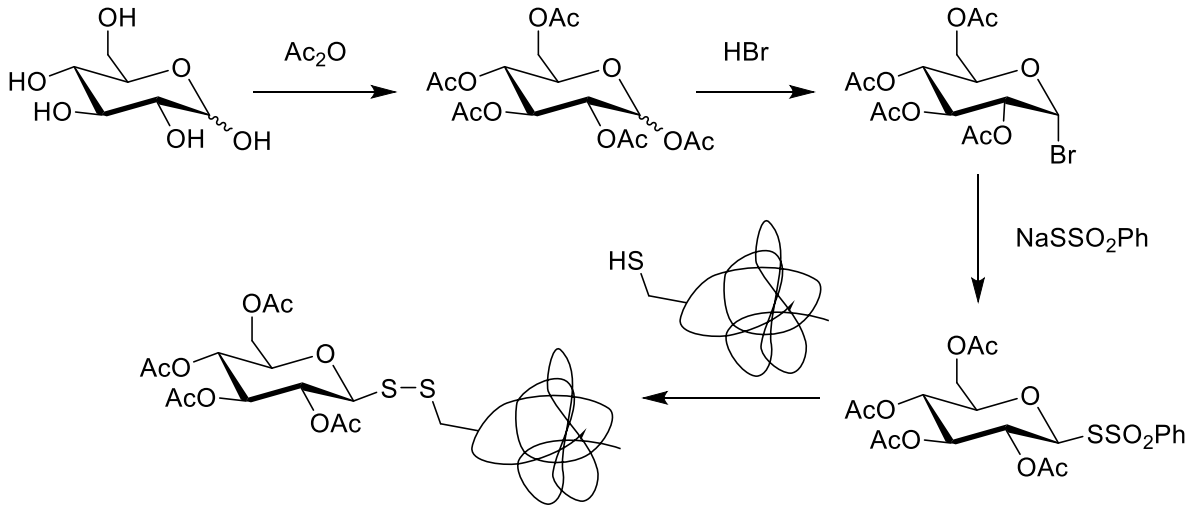
Three categories of nucleophiles have been applied to the DMC-mediated activation of sugars (**Section 1.8**): intramolecular nucleophiles, including 2-acetamido groups for the formation of oxazolines[1] and 6-hydroxyl groups for the preparation of 1,6-anhydrosugars,[2] azide for the formation of glycosyl azides,[3, 4] and thiols, mainly aryl thiols[5-7] but also alkyl thiols.[8] These nucleophiles are significantly more nucleophilic than water, and thus selective substitution at the anomeric centre is possible even when water is used as a solvent. As reported in **Chapter 2**,

mercaptoacetonitrile is another example of a nucleophile which can out-compete water, and is therefore able to effectively attack a DMC-sugar intermediate. However, cyanomethyl thioglycosides are only useful for indiscriminate glycosylation of protein lysines.[9] An investigation therefore began into DMC-mediated reactions with other nucleophiles which may allow the transformation of unprotected sugars into reagents that could subsequently be used to carry out site selective glycosylation of proteins or peptides.

Sulfur containing compounds are well-known as effective nucleophiles, and as discussed previously thiols have previously been used as nucleophiles in DMC-mediated reactions.[5, 8] Sulfur is also often used for the site-selective production of glycoconjugates, by methods which include asymmetric disulfide formation, either *via* glycosyl selenyl sulfides[10, 11] or glycosyl phenylthiosulfonates,[12] the nucleophilic addition of glycosyl thiols to DHA,[13] and the thiol-ene click chemistry of glycosyl thiols[14] (**Section 1.7.2**). However, to date the preparation of the sulfur-containing reagents required as substrates for these reactions has almost invariably required protection of the sugar.

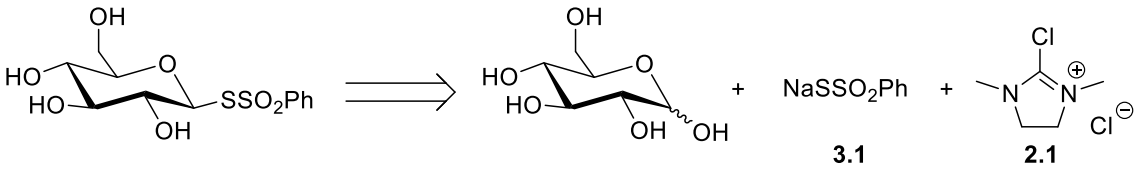
3.3 DMC-mediated preparation of glycosyl phenylthiosulfonates

Glycosyl phenylthiosulfonates have been shown to be highly effective reagents for the modification of cysteine residues of proteins.[12] However, they are prepared by reaction of sodium phenylthiosulfonate with the relevant glycosyl bromide, and therefore protection of the sugar is required (**Scheme 3.1**). Furthermore, glycosyl phenylthiosulfonates cannot be easily deprotected.



Scheme 3.1: Current preparation and application of glycosyl phenylthiosulfonates.

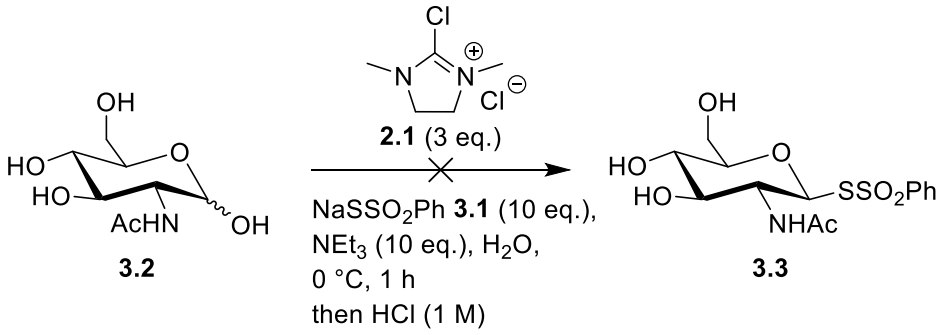
A method to prepare glycosyl phenylthiosulfonates without the use of any protecting groups would therefore greatly expand their utility. With this in mind, a DMC-mediated reaction to prepare glycosyl phenylthiosulfonates was proposed (**Scheme 3.2**).



Scheme 3.2: Retrosynthesis for the preparation of glycosyl phenylthiosulfonates using DMC.

Sodium phenylthiosulfonate **3.1** was prepared *via* the reaction of sodium benzenesulfinate with elemental sulfur in pyridine as per the previously published procedure (**Scheme 3.3**).[12]

Scheme 3.3: Preparation of sodium phenylthiosulfonate.

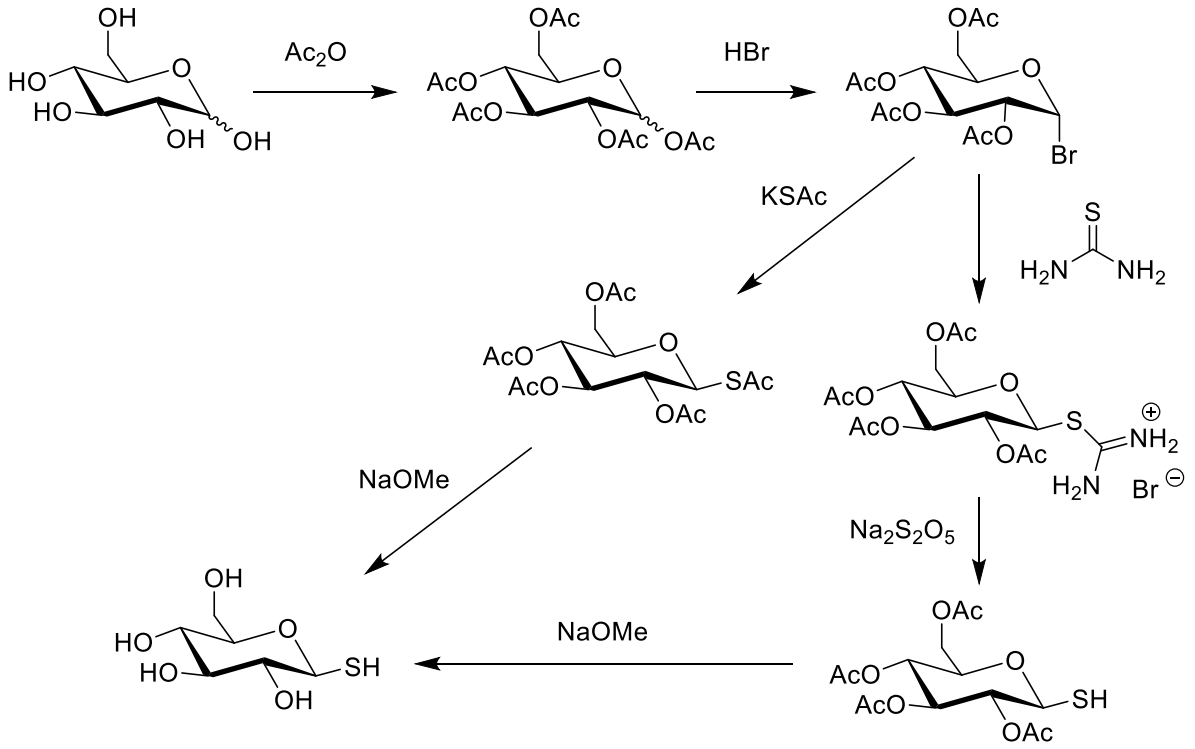


Scheme 3.4: Attempted preparation of glycosyl phenylthiosulfonates *via* DMC-mediated reaction with unprotected GlcNAc **3.2**.

Unfortunately, phenylthiosulfonate did not prove to be an effective nucleophile in DMC-mediated reactions. When GlcNAc **3.2** was used as a substrate with 3 equivalents of DMC, 10 equivalents of triethylamine as a base, and 10 equivalents of sodium phenylthiosulfonate **3.1** as a nucleophile (**Scheme 3.4**) none of the desired product **3.3** was detected by ^{1}H NMR spectroscopy of the crude reaction mixture; only starting material was observed, formed by hydrolysis of the oxazoline intermediate. The use of glucose as a staring material under similar reaction conditions resulted in the formation of 1,6-anhydroglucose, a common byproduct of the DMC-mediated activation of 2-hydroxy sugars.[2] The fact that none of the desired product was detected in either case implies that the phenylthiosulfonate anion is not a strong enough nucleophile to outcompete water when attacking the activated sugar intermediate, and therefore the synthesis of glycosyl phenylthiosulfonates *via* DMC-mediated reaction using water as a solvent seems unlikely to be successful. Furthermore, glycosyl phenylthiosulfonates have been observed to decompose when deprotected.[15] With these two points in mind, the preparation of glycosyl phenylthiosulfonates *via* a DMC-mediated reaction seemed unlikely to be successful, and so other potential nucleophiles were investigated.

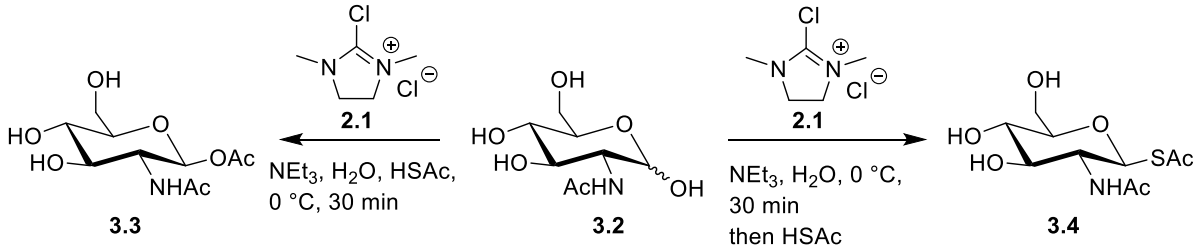
3.4 DMC-mediated preparation of glycosyl thiols

Similarly to glycosyl phenylthiosulfonates, glycosyl thiols are typically prepared by reaction of a glycosyl bromide, but with either thiourea as a nucleophile, which can be cleaved with sodium metabisulfite to afford the desired glycosyl thiol,[9] or thioacetate as a nucleophile, which can be deprotected with sodium methoxide[16] (**Scheme 3.5**). Again, protecting groups are required, and the preparation of glycosyl thiols of larger oligosaccharides is therefore challenging.



Scheme 3.5: Current preparation of glycosyl thiols.

It is also possible to prepare unprotected glycosyl thiols by direct reaction of unprotected reducing sugars with Lawesson's reagent,[11] but an apparent incompatibility with 2-acetamido groups renders this method unusable for *N*- and *O*-linked glycans, which comprise the majority of the oligosaccharides of glycoproteins.



Scheme 3.6: DMC-mediated reactions of thioacetic acid and sugars.

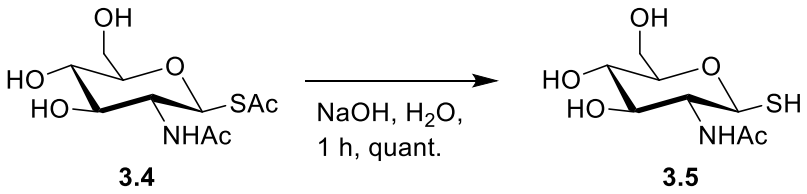
Thioacetate could be expected to attack the DMC-activated sugar intermediate (**Section 1.8**), and as was observed by another member of the Fairbanks group[17] this is indeed the case. When this reaction was carried out on GlcNAc **3.2** the product formed was glycosyl thioacetate **3.4** (**Scheme 3.6**), which could be readily deprotected to form a glycosyl thiol (**Scheme 3.5**). However, the order of reagent addition was significant, and it was found that addition of thioacetic acid at the beginning of the reaction as opposed to after 30 min resulted in the formation of the 1-*O*-acetyl sugar **3.3** instead of the 1-*S*-thioacetyl sugar **3.4**.[18]

3.5 Purification of glycosyl thioacetate 3.4

The use of 3 equivalents of DMC, 10 equivalents of triethylamine, 15 equivalents of thioacetic acid, and water as a solvent for the reaction of thioacetic acid with GlcNAc (**Scheme 3.6**) provided near-quantitative conversion to thioacetate product **3.4**. However, purification of the product was extremely challenging, as the thioacetate was both highly polar and relatively prone to hydrolysis, and at the same time a mixture of salts was present, most notably triethylammonium chloride. Purification by HPLC would be expected to provide pure product, but could only be applied on a small scale, limiting the applicability of the reaction.

Normal phase flash column chromatography did not result in adequate separation between glycosyl thioacetate **3.4** and triethylammonium chloride, and, as a result, both co-eluted using a range of solvents. Similarly, reverse phase flash column chromatography using C18 silica and HILIC flash column chromatography[19] also resulted in co-elution of **3.4** and triethylammonium chloride. In all cases separation of all other impurities was possible, but triethylammonium chloride could not be removed.

In the past Lim *et al*. solved the problem of triethylammonium salt removal by the use of a column of Amberlite® IR-120 (Na^+ form) ion exchange resin.[4] An attempt to remove the triethylammonium chloride with an Amberlite® IR-120 column resulted in partial decomposition of thioacetate and a reduced and inconsistent yield, presumably due to base-catalysed hydrolysis of the thioacetate **3.4** to form the glycosyl thiol **3.5**.



3.4 **3.5**

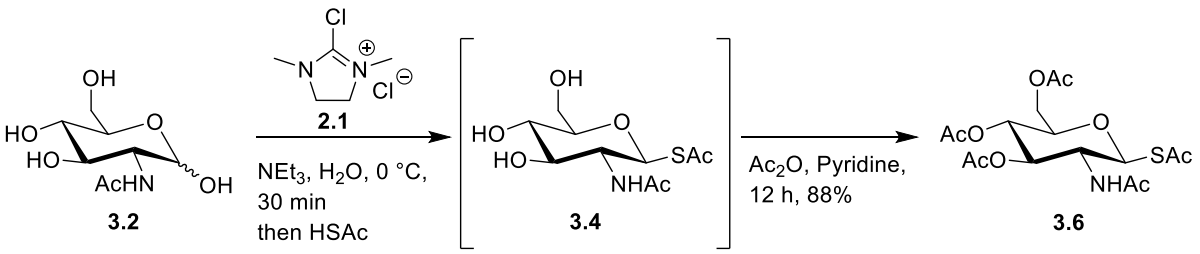
Scheme 3.7: Deprotection of thioacetate **3.4**.

Deprotection of the thioacetate using sodium hydroxide (**Scheme 3.7**) was also explored as a means of purification, as glycosyl thiol **3.5** was the ultimate goal. Treatment with a strong base was also expected to deprotonate the triethylammonium cation, which would then allow the removal of triethylamine *in vacuo*. Testing on a small quantity of purified glycosyl thioacetate **3.3** revealed that sodium hydroxide was effective for this deprotection step. Notably, a stoichiometric, rather than catalytic, quantity of base was required to deprotect the thioacetate,

as the glycosyl thiolate produced by the reaction is not sufficiently basic to deprotonate water and thus continue the catalytic cycle.

Treatment of the crude glycosyl thioacetate **3.4** with sodium hydroxide did facilitate the removal of triethylammonium chloride. However, significant quantities of sodium acetate, sodium thioacetate, and sodium chloride remained as impurities. Flash column chromatography was not effective for removal of these impurities, as glycosyl thiol **3.5** displayed a tendency to bind strongly to silica, which resulted in a low isolated yield of product.

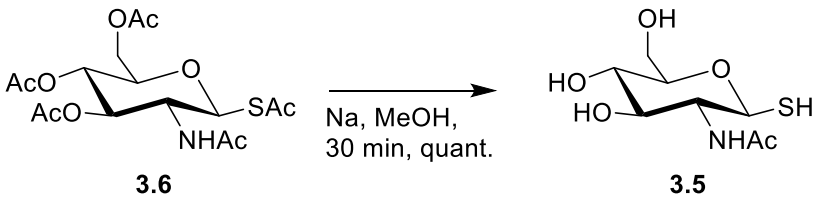
Investigation of different workup procedures to remove all the salt impurities met with more success. Following the reaction, extraction with dichloromethane followed by trituration with ethyl acetate removed all by-products except triethylammonium chloride. Subsequent deacetylation with sodium methoxide and trituration with methanol then afforded clean glycosyl thiol **3.5**.



Scheme 3.8: Purification of glycosyl thioacetate **3.4** *via* acetylation.

As the method detailed above is rather involved, and does not include a chromatography or recrystallization step and therefore may not always provide high purity product, an alternative method was also investigated. Full protection of the sugar significantly increased the ease of purification, and could be carried out in one pot. Protection can be avoided if necessary, for example in the preparation of glycosyl thiols where only a small amount of sugar is available

and minimising the number of steps is therefore desirable in order to maximise the yield, by purification of the glycosyl thioacetate by HPLC. Crude glycosyl thioacetate **3.4** was acetylated with acetic anhydride in pyridine[20] to afford acetylated glycosyl thioacetate **3.6** in a pleasingly high yield of 88% from GlcNAc. Purification was then easily accomplished by normal phase flash column chromatography.



3.6 **3.5**

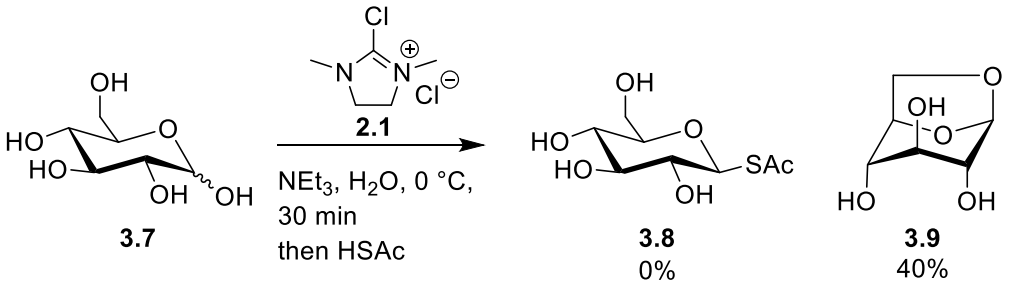
Scheme 3.9: Deacetylation of peracetylated glycosyl thioacetate **3.6**.

Finally, deprotection was easily accomplished in quantitative yield by treatment of acetylated glycosyl thioacetate **3.6** with sodium in methanol (**Scheme 3.9**),[21] giving an overall yield of 88% for the preparation of glycosyl thiol **3.5** from GlcNAc **3.2**.

3.6 Attempted extension of the DMC-mediated reaction of sugars with thioacetic acid to 2-hydroxy sugars

The most notable difference between 2-acetamido and 2-hydroxy sugars in DMC-mediated reactions is that while 2-acetamido sugars react *via* a relatively stable 1,2-oxazoline intermediate,[1] the intermediate in the reaction of 2-hydroxy sugars is very unstable, and has never been conclusively observed (**Section 1.8**). This leads to a number of practical differences in the application of DMC activation to the two varieties of sugars. Firstly, when the 2-hydroxy group is equatorial the corresponding 1,6-anhydrosugar is invariably formed as a by-product. Secondly, because the intermediate in the reaction of 2-hydroxy sugars is significantly less

stable than that formed during the reaction of 2-acetamido sugars, the addition of an acid to open the cyclic intermediate[4, 6] is not required.



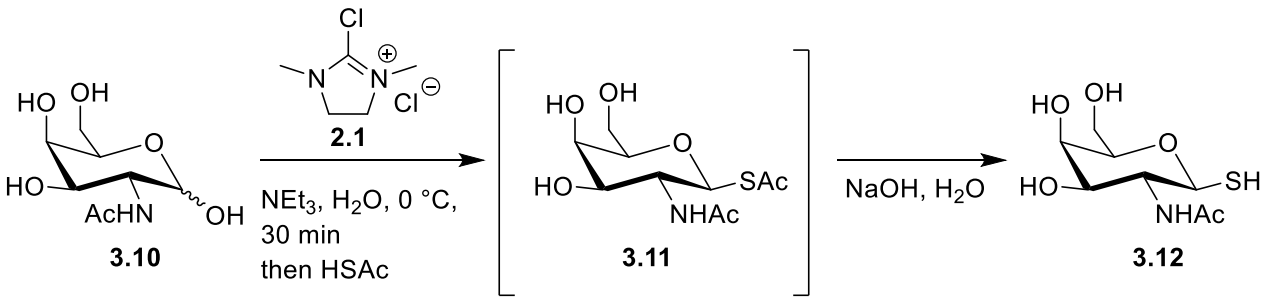
Scheme 3.10: DMC-mediated reaction of thioacetic acid and glucose.

Unfortunately, the reaction of glucose **3.7** under the conditions employed for GlcNAc **3.2** (**Scheme 3.6**) resulted in the formation of no thioacetate **3.8**; rather, 1,6-anhydroglucose **3.9** was isolated in a 40% yield, with the remainder of the material being unreacted glucose **3.7** (**Scheme 3.10**). Increasing or decreasing the activation time between the addition of DMC and the addition of the thioacetic acid did not lead to the formation of any product, presumably because upon the addition of thioacetic acid no further sugar activation by DMC occurs due to the then acidic conditions. Under basic conditions DMC reacts preferentially with thioacetate rather than glucose **3.7** to form a DMC-thioacetate intermediate, which then reacts with sugars to produce glycosyl acetates.[18]

As a consequence of these results, investigation into the formation of anomeric thioacetates of 2-hydroxy sugars was halted, as it was judged to be unlikely to be successful. However, this is only a minor drawback of this method of preparing glycosyl thioacetates for use in conjugation reactions, because, as discussed previously, the majority of natural glycans have 2-acetamido sugars at the reducing terminus.

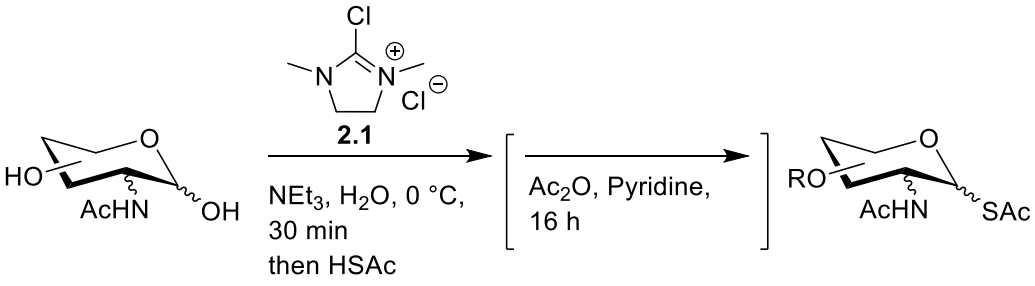
3.7 Conversion of a range of 2-acetamido sugars into their glycosyl thioacetates

With optimised conditions in hand, the DMC-thioacetic acid reaction was tested on a range of 2-acetamido sugars.



Scheme 3.11: Preparation of the glycosyl thiol of *N*-acetylgalactosamine.

N-Acetylgalactosamine (GalNAc) thiol **3.12** was prepared from GalNAc **3.10** by a similar procedure to that used to make GlcNAc thiol **3.5** (**Scheme 3.11**) and purification was attempted *via* a workup, as previously (**Section 3.5**). Unfortunately, this did not result in clean thiol **3.12**, meaning that the sequence of washes that had been developed for the synthesis of GlcNAc thiol **3.5** is not generally applicable, and implying that a new procedure would have to be developed for each sugar examined. Furthermore, the reaction of *N*-acetylgalactosamine to form thioacetate **3.11** did not go to completion under the standard reaction conditions, and a number of minor sugar impurities were present. The remaining starting material and other sugar impurities were not removed by the workup procedure, exposing one of the key limitations of the workup method. The reaction was therefore exemplified with monosaccharides *via* the protection/purification/deprotection strategy outlined in **Schemes 3.8** and **3.9**, whilst larger oligosaccharides were purified by HPLC.



Scheme 3.12: Preparation of glycosyl thioacetates.

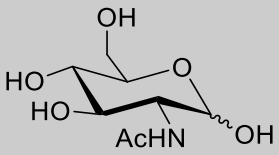
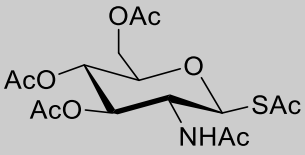
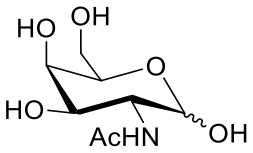
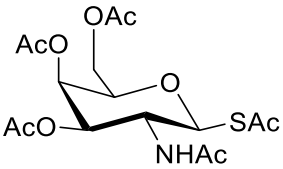
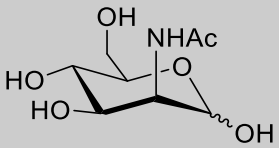
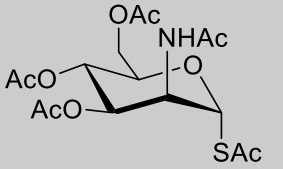
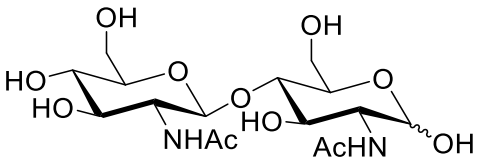
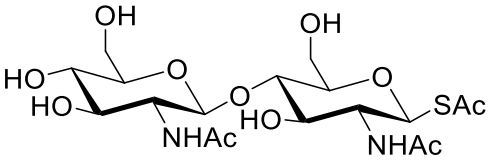
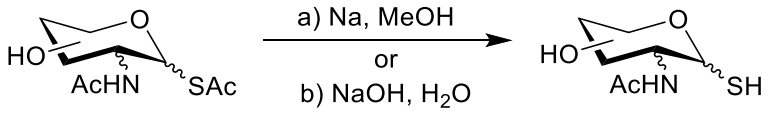
Sugar	Product	Yield
	 **3.6**	88%[a]
	 **3.13**	56%[a]
	 **3.14**	89%[a]
	 **3.15**	93%[b]

Table 3.1: Preparation of glycosyl thioacetates. [a]Purified by acetylation followed by flash column chromatography. [b]Purified by HPLC.

Pleasingly, high yields of the glycosyl thioacetate were obtained in all cases except for GalNAc thioacetate **3.13**. As discussed in **Section 2.4**, activation of galactosyl sugars by DMC results in relatively low conversion to product and the formation of a number of unidentified byproducts, so this result was not unexpected. The glycosyl thioacetate of *N,N'*-diacetylchitobiose (**3.15**) was formed in high yield without difficulty, confirming that the reaction is applicable to sugars larger than a monosaccharide.

The selection of glycosyl thioacetates were then deprotected using previously established conditions (**Scheme 3.13**). Thioacetates **3.6**, **3.13**, and **3.15** were deprotected in quantitative yield using either sodium in methanol or sodium hydroxide in water (**Table 3.2**).



Scheme 3.13: Preparation of glycosyl thiols.

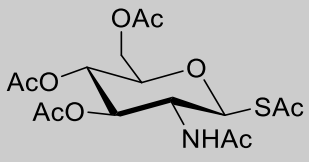
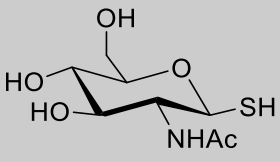
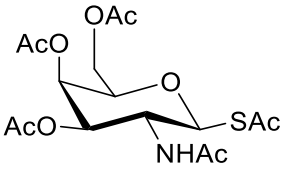
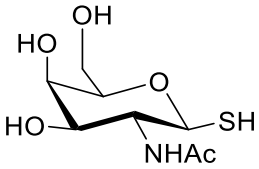
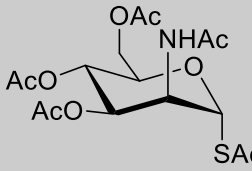
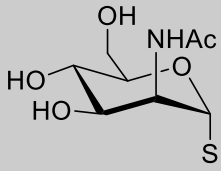
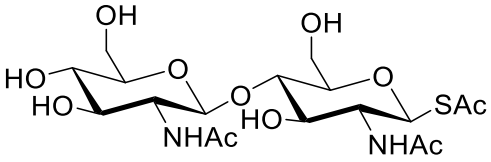
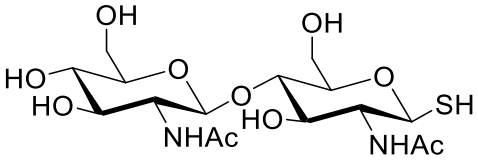
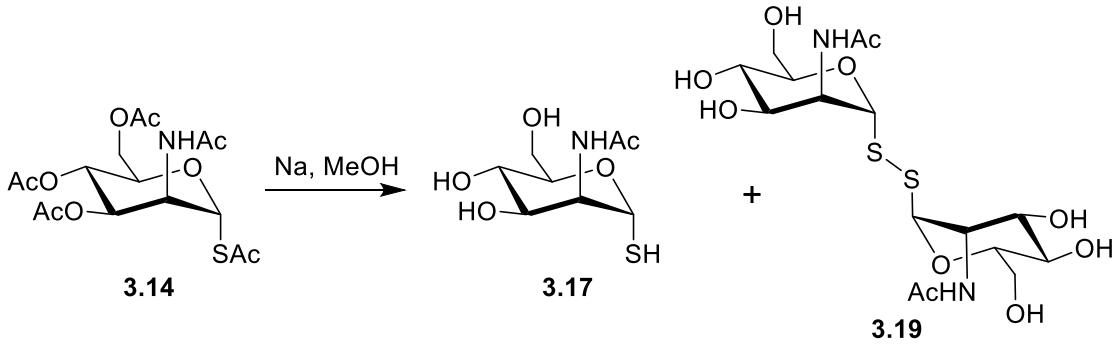
Starting Material	Product	Yield
 **3.6**	 **3.5**	quant.[a]
 **3.13**	 **3.16**	quant.[a]
 **3.14**	 **3.17**	decomp.[a]
 **3.15**	 **3.18**	quant.[b]

Table 3.2: Preparation of glycosyl thiols. [a]Method a). [b]Method b).

Unfortunately, *N*-acetylmannosamine (ManNAc) thioacetate **3.14** was not readily deprotected under these conditions. While the acetyl groups were indeed removed, a number of other products were formed, including disulfide **3.19** (**Scheme 3.14**). Attempts to carry out the deacetylation reaction in the presence of reducing agents, including dithiothreitol and

TCEP,[22, 23] did not significantly improve the conversion to product, and as established previously the purification of unprotected glycosyl thiols is challenging so separation of glycosyl thiol **3.17** from disulfide **3.19** is not possible. ManNAc thiol **3.17** was not an attractive target: the primary biological role of ManNAc is the biosynthesis of *N*-acetylneuraminic acid,[24] and it is not found in natural glycoprotein structures. Attempts to optimise the reaction for ManNAc were therefore abandoned.



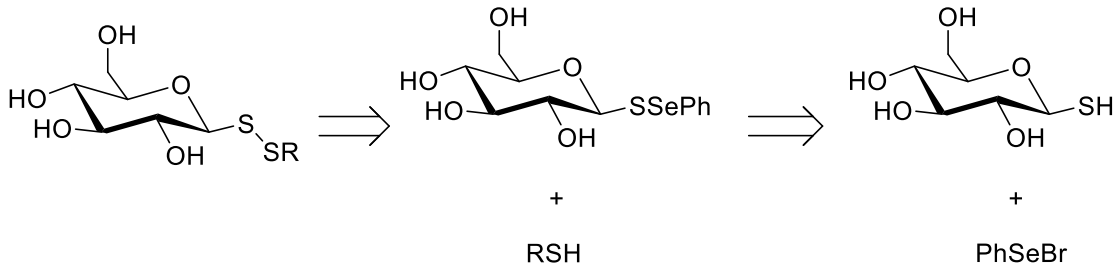
Scheme 3.14: Deprotection and oxidation of ManNAc thiol **3.14**.

3.8 Preparation of glycoconjugates

A number of options exist for the site-selective preparation of glycoconjugates from glycosyl thiols. These include the production of glycoproteins and glycopeptides by the attachment of glycosyl thiols to cysteine residues through the use of selenyl sulfides,[10] the reaction of glycosyl thiols with DHA,[13] and the use of photocatalysed thiol-ene click chemistry to attach glycosyl thiols to alkenes.[14] Of these methods, selenyl sulfide chemistry was selected for initial exemplification, due to the apparent ease of applying it to larger structures, including proteins.

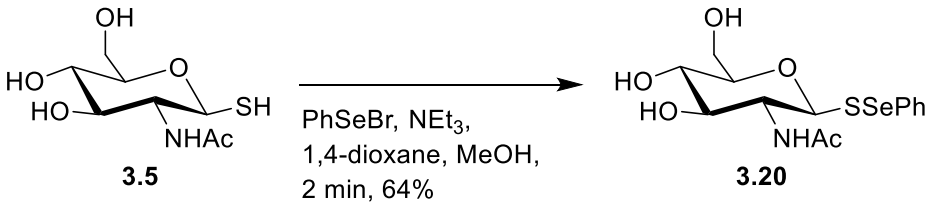
3.8.1 Preparation of glycoconjugates *via* selenyl sulfide chemistry

As discussed in **Section 1.7.2**, glycosyl selenyl sulfides are effective reagents for the modification of thiols, including the cysteine residues of proteins.



Scheme 3.15: Use of glycosyl selenyl sulfides to prepare glycoconjugates.

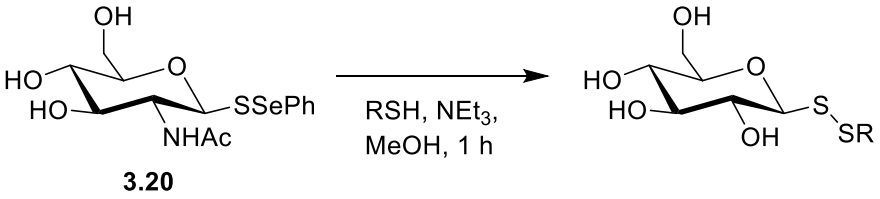
In order to prepare glycoconjugates using glycosyl selenyl sulfides, one of the thiols is first reacted with phenylselenyl bromide, and the resulting selenyl sulfide is reacted with the desired thiol to produce a glycoconjugate (**Scheme 3.15**).[10]



Scheme 3.16: Preparation of GlcNAc selenyl sulfide **3.20**.

The glycosyl selenyl sulfide **3.20** of GlcNAc thiol **3.5** was prepared in 64% yield following the literature method (**Scheme 3.16**).[10] The dropwise addition of a solution of **3.5** to a solution of phenylselenyl bromide proved to be critical to obtaining even a moderate yield of product. Any other order of addition resulted in the formation of a very small amount of product due to formation of disulfide by reaction between glycosyl thiol **3.5** and product **3.20**.

Two asymmetric disulfides were then prepared using glycosyl selenyl sulfide **3.20** as one component in order to test the selenyl sulfide methodology (**Scheme 3.17**).



3.20

Scheme 3.17: Reaction of selenyl sulfide **3.20** with thiols.

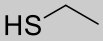
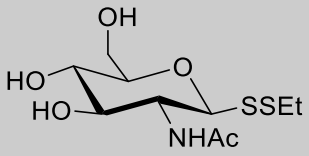
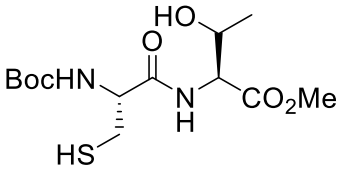
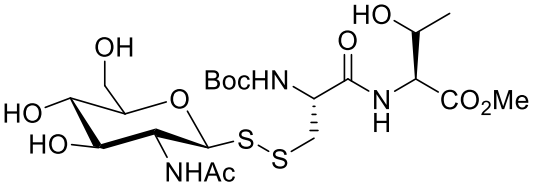
Thiol	Product	Yield
	 **3.21**	95%
 **3.22**	 **3.23**	77%

Table 3.3: Reactions of selenyl sulfide **3.20** with thiols.

Selenyl sulfide **3.20** reacted with ethanethiol and dipeptide **3.22** to give the corresponding disulfide products **3.21** and **3.23** in yields of 95% and 77% respectively. In both cases 3 equivalents of sugar **3.20** were used, and therefore the yield was calculated with respect to the thiol.

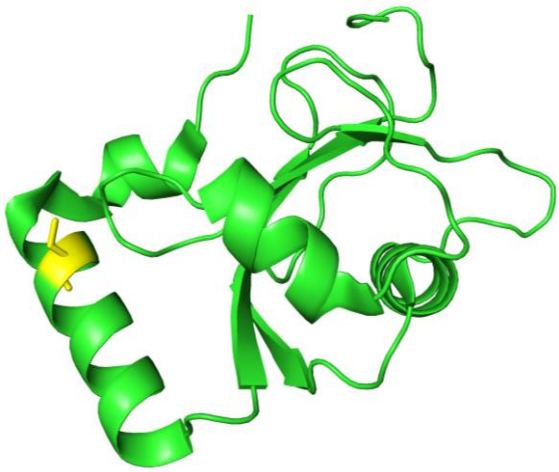
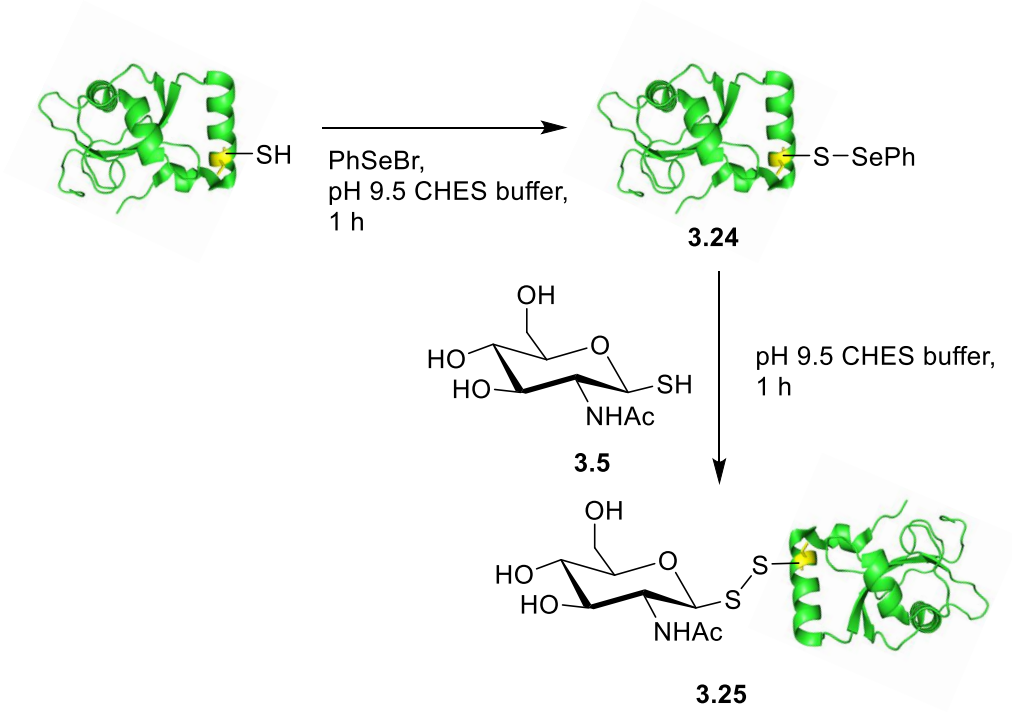


Figure 3.1: Crystal structure of GABARAPL2, with Cys15 highlighted in yellow.

Having demonstrated that glycosyl selenyl sulfide reactions could successfully be carried out with a glycosyl thiol prepared *via* the DMC activation of a free sugar, the next goal was to attach GlcNAc to a protein *via* selenyl sulfide chemistry. Gamma-aminobutyric acid receptor-associated protein-like 2 (GABARAPL2) which was available in the University of Canterbury Biochemistry Department, was selected for exemplification. GABARAPL2 has a single cysteine (**Figure 3.1**, Cys15, highlighted in yellow),[25] and was expressed by Dr. Akshita Wason. BSA also contains a single free cysteine residue, but was not used as a test protein as previous studies with glycosyl selenyl sulfides showed that modification of the single cysteine of BSA not involved in a disulfide bond was challenging.[10]



Scheme 3.18: Preparation of a GABARAPL2 glycoprotein.

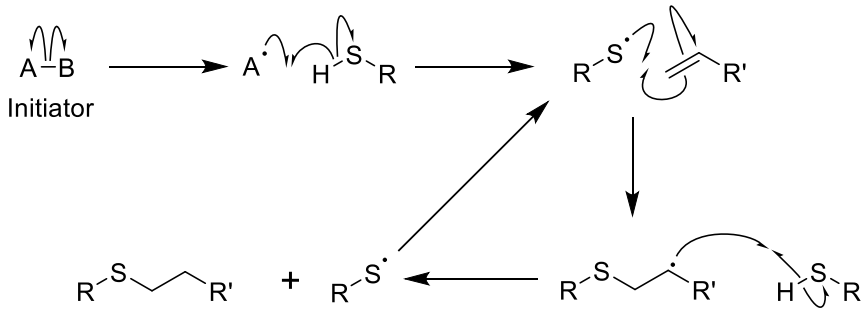
Modification of GABARAPL2 was carried out using the literature conditions (**Scheme 3.18**).[10] First, selenyl sulfide protein **3.24** was prepared by reaction of GABARAPL2 with phenylselenyl bromide in pH 9.5 CHES buffer, and then glycoprotein **3.25** was prepared by reaction of glycosyl thiol **3.5** with selenyl sulfide protein **3.24**. While this reaction proved successful in the first instance, the results varied considerably between trials. In many cases neither product nor starting material was observed upon MS analysis. Variation of buffer, purification technique, and MS conditions did not provide greater consistency. It appeared that GABARAPL2 shows a tendency to aggregate, and for this reason tended to produce inconsistent results. Conditions to avoid this aggregation could not be found.

As no suitable protein was available for the application of glycosyl thiols to the selenyl sulfide reaction, and furthermore the preparation of glycosyl selenyl sulfides was found to be

inconsistent, a different method for the preparation of glycoconjugates using glycosyl thiols was investigated.

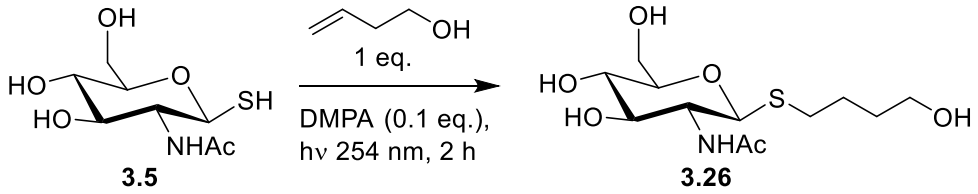
3.8.2 Preparation of glycoconjugates *via* thiol-ene click chemistry

Photocatalysed thiol-ene click chemistry is a versatile reaction for the coupling of alkenes and thiols (**Scheme 3.19**).[14] It can be performed with a wide variety of thiols and alkenes, in a wide variety of solvents, and provides thioether-linked products in high yields and with good selectivity. Alkenes are not typically found as part of proteins or peptides, and so represent an attractive functional group for the site-selective modification of these species, as has previously been demonstrated (**Section 1.7.2**).[26]



Scheme 3.19: Mechanism of the photocatalysed thiol-ene click reaction.

With the aim of applying the thiol-ene click reaction to the preparation of glycoconjugates, suitable initial conditions were investigated. 2,2-Dimethoxy-2-phenylacetophenone (DMPA) was selected as the initiator,[27] as it has previously been applied to the thiol-ene click reaction between sugars and proteins,[28] and is readily available. Water was selected as a solvent, as it is both benign and readily dissolves unprotected glycosyl thiols. 3-Buten-1-ol was chosen as a test alkene, as it is water-soluble, simple, and readily available.



Scheme 3.20: Thiol-ene click reaction of glycosyl thiol **3.5**.

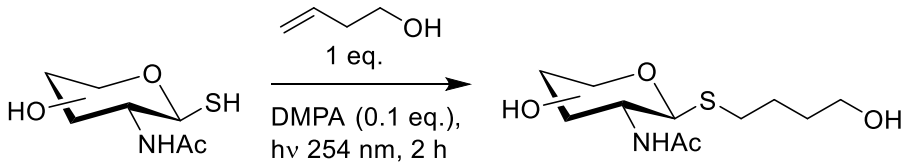
Experiment	DMPA Equivalents	Conversion	Isolated Yield
Standard	0.1	89%	61%
Control	0	8%	-

Table 3.4: Photocatalysed thiol-ene click reaction of **3.5**.

The photocatalysed thiol-ene click reaction proved to be highly effective when applied to GlcNAc thiol **3.5** (**Scheme 3.20**). Thioether **3.26** was obtained with a conversion to product (as determined by ^{1}H NMR spectroscopy) of 89% and an isolated yield of 61% (**Table 3.4**); the lower isolated yield reflects the difficulty of purifying amphiphilic **3.26** by flash column chromatography. A control reaction in which no DMPA was added resulted in only minor conversion to product, confirming that the thiol-ene reaction requires a radical initiator in order to proceed.

3.8.2.1 Exemplification: variation of glycosyl thiol

With confirmation that the thiol-ene click reaction worked extremely effectively when a glycosyl thiol and 3-buten-1-ol were used as substrates, it was extended to the other glycosyl thiols that had been prepared previously (**Section 3.5**). Yields of glycoconjugate product between 60% and 65% were obtained for all sugar substrates (**Scheme 3.21**, **Table 3.5**), although, as for **3.26**, the conversion to product was significantly higher in all cases.



Scheme 3.21: Thiol-ene reaction of glycosyl thiols.

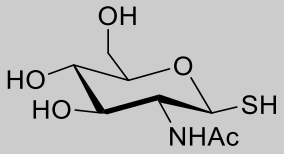
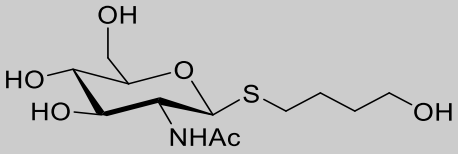
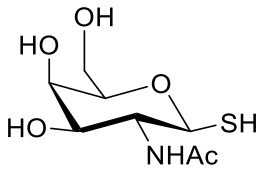
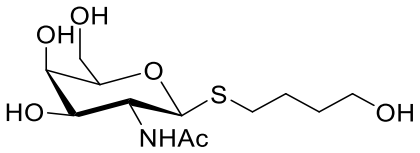
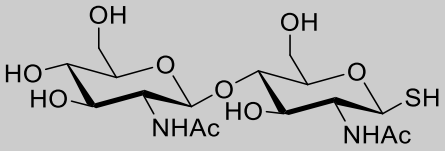
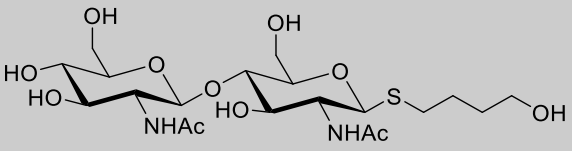
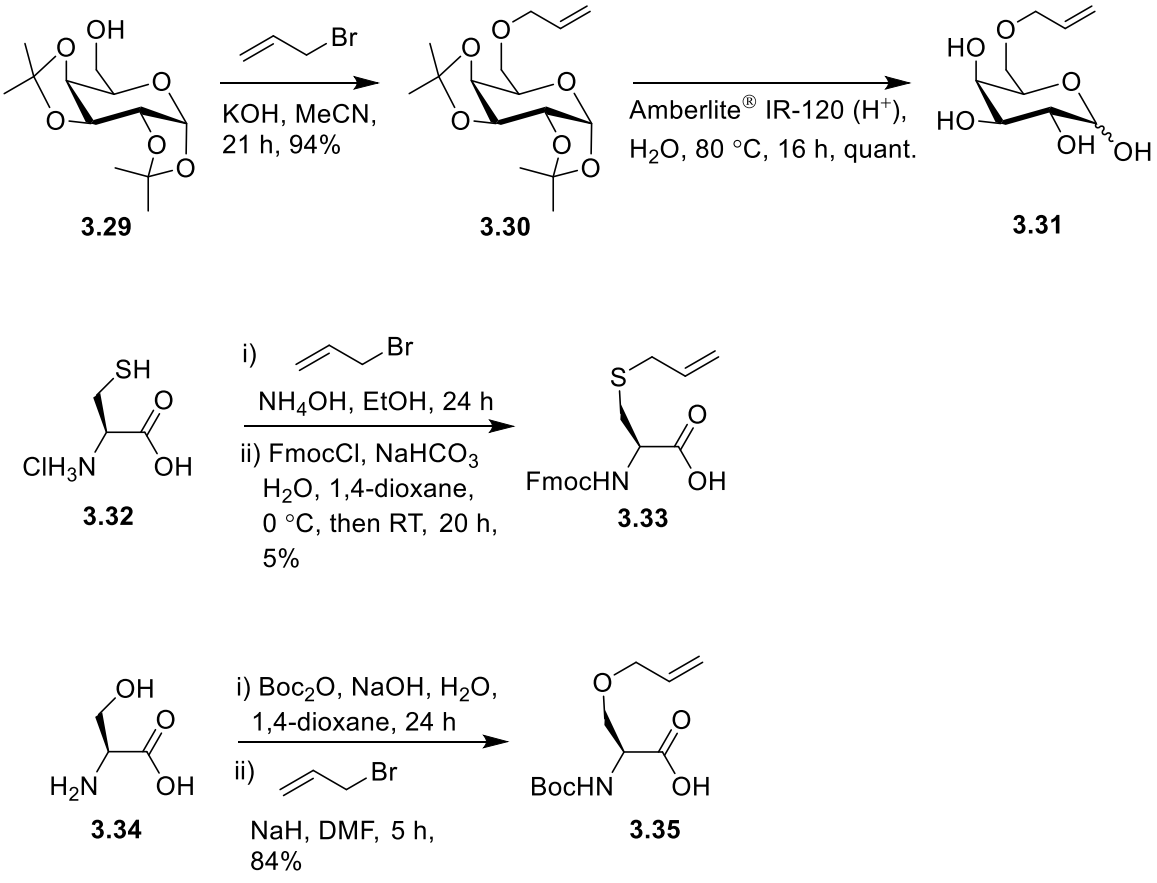
Glycosyl Thiol	Product	Yield
 **3.5**	 **3.26**	61%
 **3.16**	 **3.27**	61%
 **3.18**	 **3.28**	64%

Table 3.5: Thiol-ene click reaction of 3-buten-1-ol and glycosyl thiols **3.5**, **3.16**, and **3.18**.

3.8.2.2 Exemplification: variation of alkene

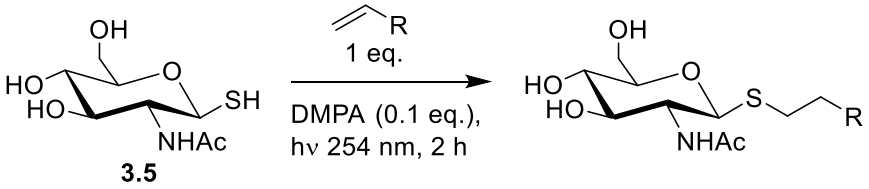
As the glycosyl thiol substrate could demonstrably be changed without altering the efficiency of the thiol-ene click reaction, variation in the alkene substrate was then examined.

In order to carry out this investigation, a selection of biologically relevant alkenes were prepared (**Scheme 3.22**). 6-*O*-Allyl galactose **3.31**, to be used for the preparation of a disaccharide mimic, was prepared in two steps from commercially available diacetone galactose **3.29**, by allylation at position 6 using potassium hydroxide and allyl bromide to afford protected allylated galactose **3.30**, followed by deprotection of the two acetonides under acidic conditions to afford unprotected allyl galactose **3.31**.[29]



Scheme 3.22: Preparation of alkene substrates.

Two protected amino acid alkenes based on cysteine **3.32**[30] and serine **3.34**[31] were also prepared to demonstrate the use of glycosyl thiols in preparing glycopeptides. *N*-Protection is important, as it would allow the use of the conjugated amino acids in subsequent solid phase peptide synthesis. Fmoc-Protected allyl cysteine **3.33** was prepared by allylation of cysteine hydrochloride **3.32** using ammonium hydroxide as a base, followed by Fmoc protection using Fmoc chloride, with sodium bicarbonate as a base. Boc-Protected allyl serine **3.35** was prepared by Boc protection of serine **3.34** using Boc anhydride and sodium hydroxide, followed by allylation with allyl bromide and sodium hydride.



Scheme 3.23: Testing of the alkene substrate scope of the thiol-ene click reaction.

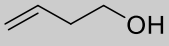
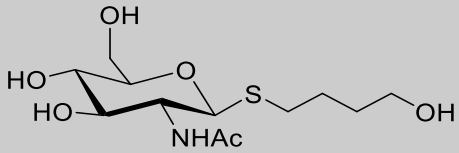
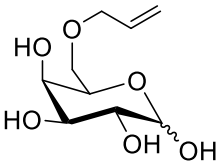
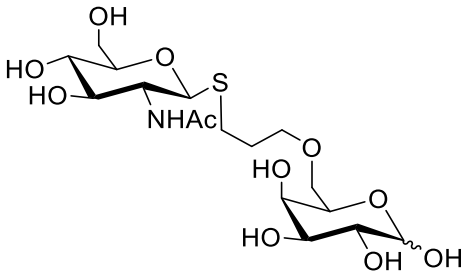
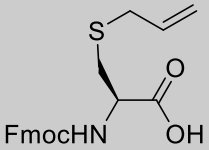
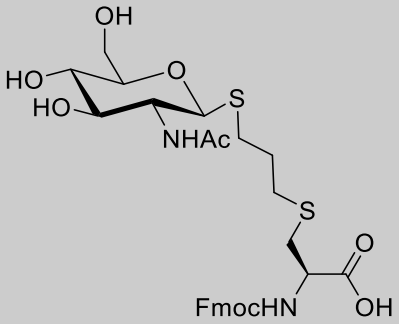
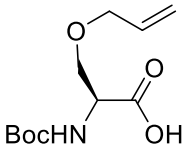
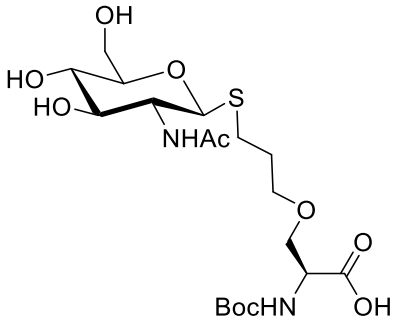
Alkene	Product	Solvent	Yield
	 **3.26**	H_2O	61%
 **3.31**	 **3.36**	H_2O	73%
 **3.33**	 **3.37**	MeOH	decomp.
 **3.35**	 **3.38**	MeOH:H_2O 1:1	91%

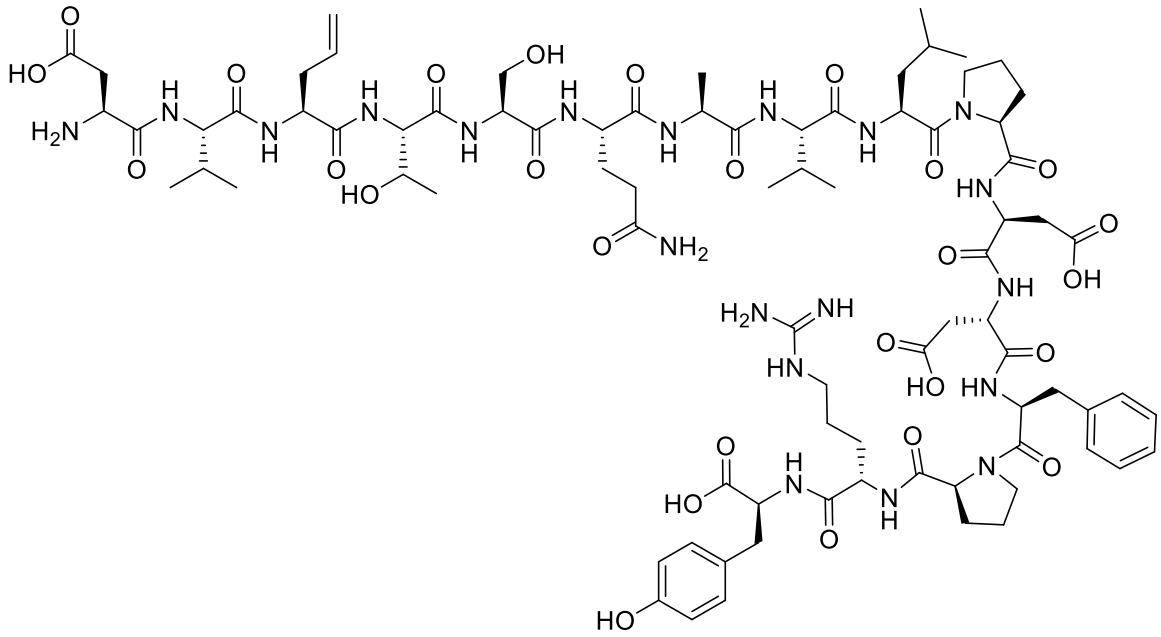
Table 3.6: Variation of the alkene in the thiol-ene click reaction.

With this selection of alkenes available, testing of the thiol-ene click reaction commenced (**Scheme 3.23**). For 6-*O*-allyl galactose **3.31** the reaction proved effective (**Table 3.6**), and gave the disaccharide mimic **3.36** in 73% yield. Unfortunately, allyl cysteine **3.33** did not prove suitable for the thiol-ene click reaction, with a number of decomposition products observed in the crude reaction mixture, a result consistent with a previous study by Fiore *et al.*,[32] who reported that the use of an allyl linker attached to cysteine led to the formation of a number of by-products derived from the coupling of various thiyl radicals. However, allyl serine **3.35** produced glycosyl amino acid **3.38** in an impressively high yield of 91%, demonstrating the potential that thiol-ene click chemistry has for the preparation of glycopeptides and glycoproteins.

The thiol-ene click reaction has been shown to be an effective reaction for the formation of simple glycoconjugates. However, most interesting natural glycoconjugates are significantly larger than those examined up to this point in this thesis, and therefore attention turned to the preparation of a more complex, biologically relevant glycopeptide.

3.9 Application of the thiol-ene click reaction to a model biological system

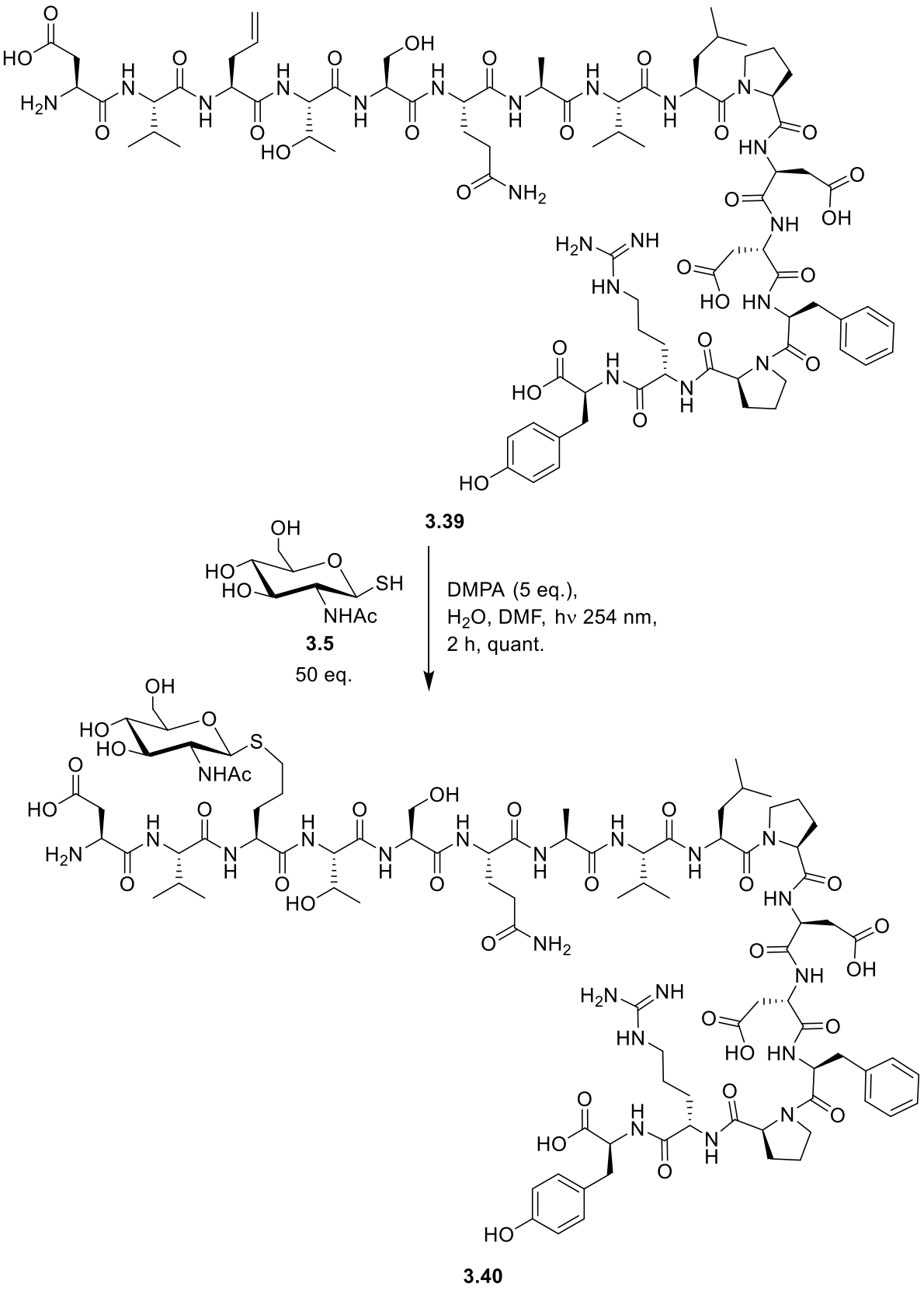
While insertion of an alkene into a protein is complex,[26] insertion of an alkene into a synthetic peptide is relatively simple. With this in mind, the next goal was the attachment of a glycosyl thiol to a peptide. Peptide **3.39**, with a single allylglycine residue at position 3 (**Figure 3.2**), was prepared and generously provided by Zaid Anso of the Brimble group. The peptide selected, a section of the rat pancreatic hormone preptin,[33] was a readily available peptide with a single allylglycine residue, making it suitable for use as a model peptide.



3.39

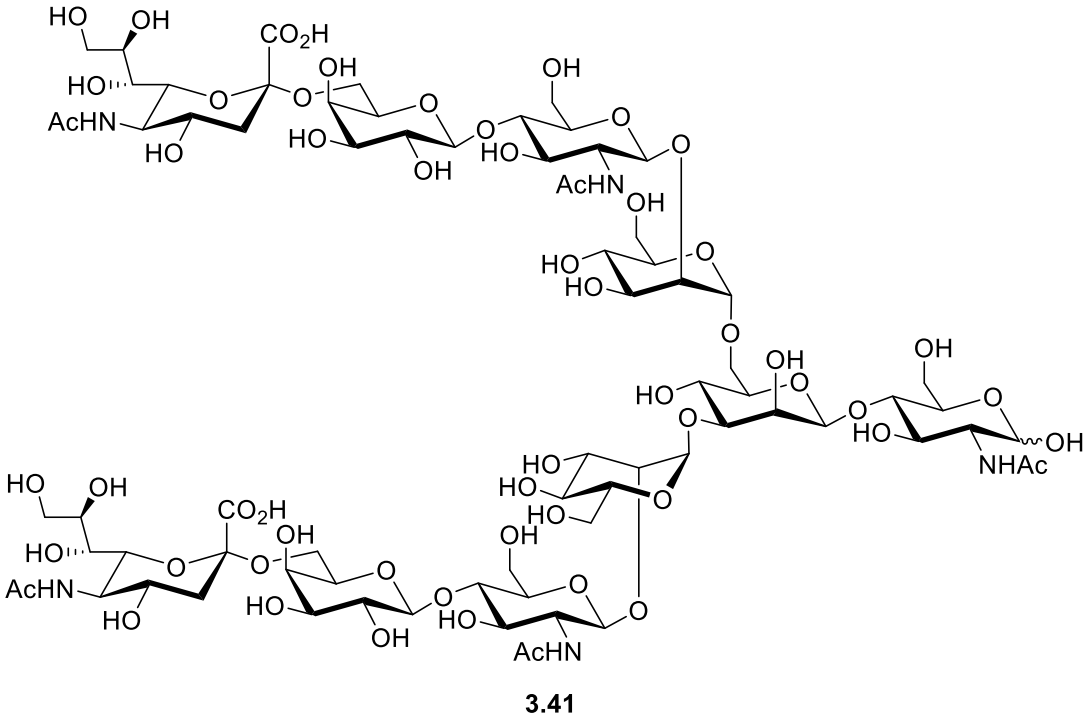
Figure 3.2: Modified preptin **3.39** containing a single allylglycine residue for use in the thiol-ene click reaction.

Firstly, the thiol-ene reaction of GlcNAc thiol **3.5** with peptide **3.39** was examined. Pleasingly, the use of 5 equivalents of DMPA in water with a large excess of thiol **3.5** and irradiation at 254 nm resulted in quantitative conversion to product glycopeptide **3.40** as assessed by HPLC (**Scheme 3.24**). The product was also characterised by LC-MS, confirming the complete conversion of peptide **3.39** to glycopeptide **3.40**.



Scheme 3.24: Preparation of glycopeptide **3.40** *via* thiol-ene click reaction.

While attachment of a monosaccharide demonstrated that the thiol-ene click reaction was effective for larger alkene substrates, most glycopeptides and glycoproteins also contain large oligosaccharides. The next step was therefore to prepare and attach a large oligosaccharide to peptide **3.39**.

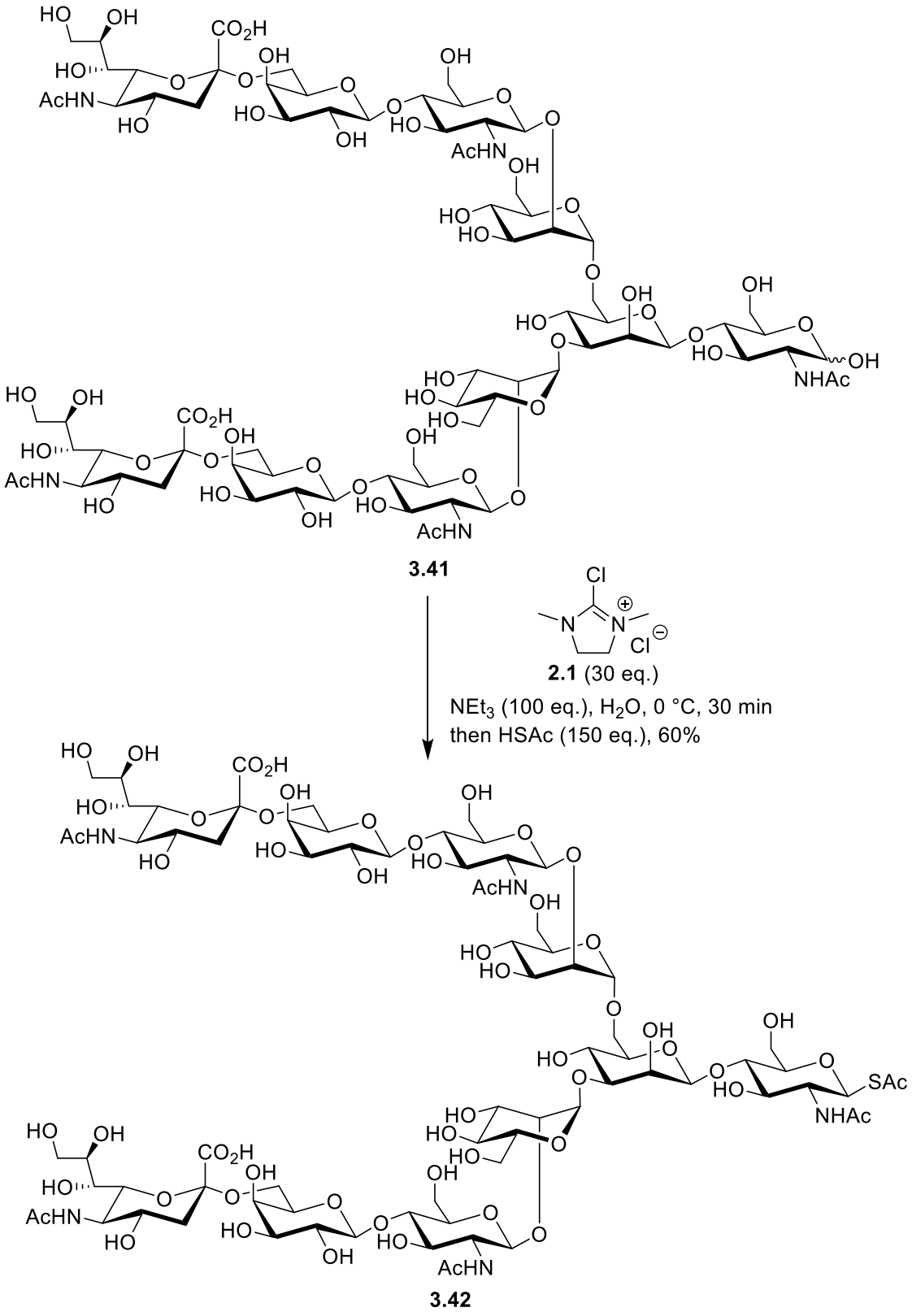


3.41

Figure 3.3: Complex biantennary glycan isolated from hens' egg yolks.

Sialoglycan **3.41** (**Figure 3.3**) is readily isolated from hens' egg yolks,[34, 35] and therefore represents a suitable oligosaccharide for exemplification of the methodology developed in this chapter. The sample of sialoglycan **3.41** used in this study was prepared by Dr. Vivek Poonthiyil.

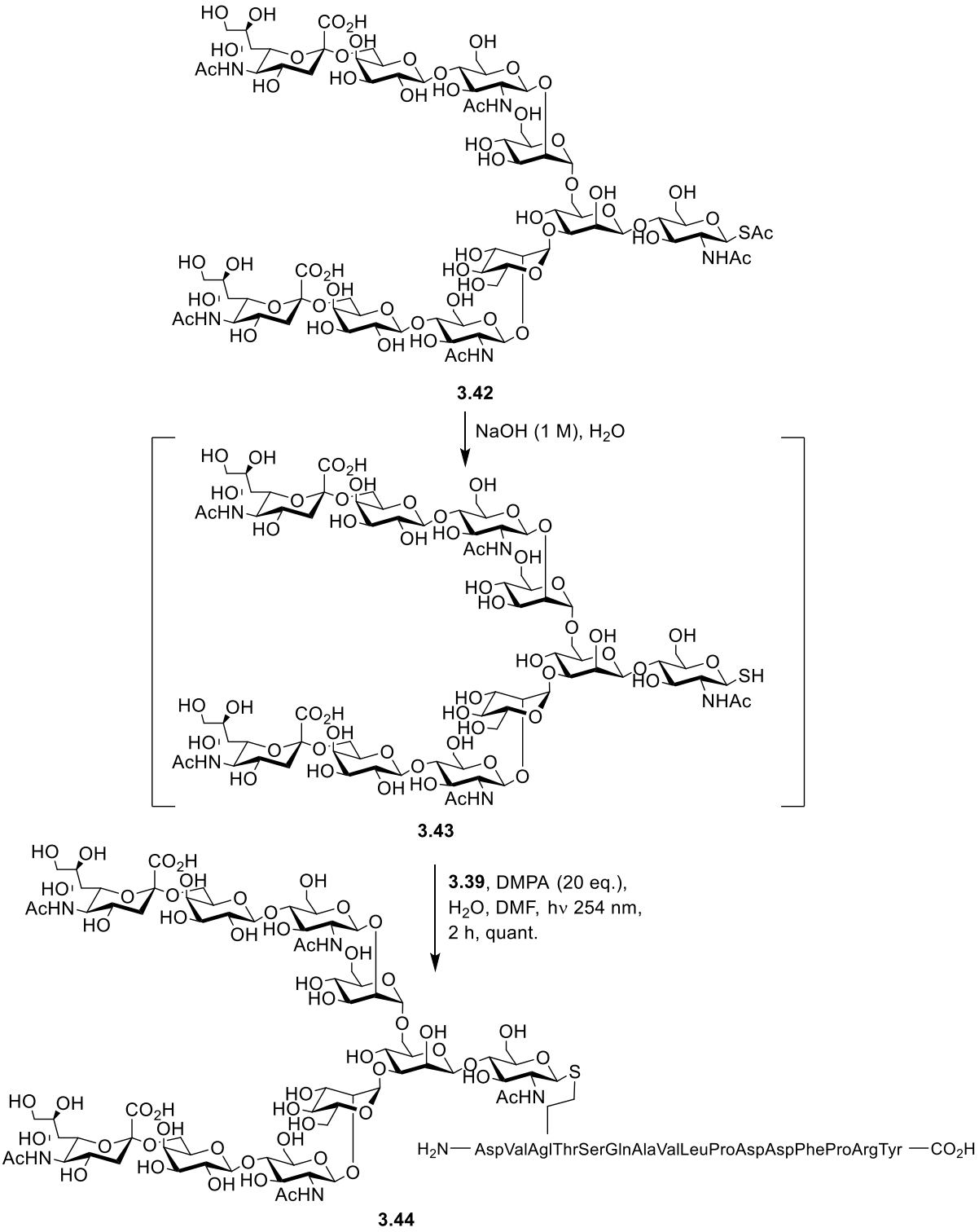
The DMC-mediated reaction between thioacetic acid and sialoglycan **3.41** proved successful (**Scheme 3.25**). However a greater number of equivalents of DMC, triethylamine, and thioacetic acid were required.



Scheme 3.25: Preparation of sialoglycan glycosyl thioacetate **3.42**.

The reaction appeared to be less efficient on sialoglycan **3.41** than on mono- and disaccharides, and this was partially attributed to the fact that the rate of addition of both DMC and thioacetic acid to the reaction mixture required careful control: rapid addition at either of these stages typically led to an incomplete reaction. Such control of the rate of addition was practically challenging for the addition of very small volumes, and therefore the use of a greater quantity of reagents led to a more efficient reaction. As all the reagents were readily available, increasing the quantities did not present a significant hurdle. Using this method, glycosyl thioacetate **3.42** was obtained in 60% yield.

The next step was to deprotect sialoglycan thioacetate **3.42** and then attach it to peptide **3.39**. Again, the reaction proved straightforward (**Scheme 3.26**). Thioacetate **3.42** was de-protected with aqueous sodium hydroxide to afford glycosyl thiol **3.43**, which was then reacted with peptide **3.39** without further purification to afford glycopeptide **3.44** in quantitative yield, as determined by HPLC. A relatively large excess of thiol **3.43** (800 equivalents) was required in order to ensure complete conversion: the use of a reduced number of equivalents of thiol **3.43** resulted in some conversion to glycopeptide, but a significant amount of unreacted peptide **3.39** remained, as observed by mass spectrometry. This is likely to be due to disulfide formation, a side reaction caused by the combination of two thiyl radicals, which is more likely to have a significant effect on the reaction when the relative rate of the reaction between the thiyl radical and alkene is slow, as would be expected for the two large, sterically hindered species examined here. Nevertheless, it has been demonstrated that thiol-ene click chemistry is effective for the preparation of complex glycopeptides.



Scheme 3.26: Application of the thiol-ene click reaction to the preparation of a complex glycopeptide.

3.10 Conclusions

DMC was applied to the preparation of glycosyl thiols in aqueous conditions without the use of protecting groups. The use of 3 equivalents of DMC, 10 equivalents of triethylamine, and 15 equivalents of thioacetic acid proved effective in most cases for 2-acetamido sugars, while 2-hydroxy sugars proved unsuitable for the reaction. Furthermore, the glycosyl thiols prepared *via* this method were then used to prepare glycoconjugates, culminating in the attachment of a decasaccharide to a 16mer peptide, demonstrating that the methodology developed in this chapter provides a simple, robust means of site-selectively modifying glycopeptides.

Future studies could include the attachment of glycosyl thiols to a protein *via* thiol-ene click chemistry, and the preparation of more biologically relevant glycopeptides with a view towards testing any impact of the unnatural thioether linker on their biological activity.

3.11 References

1. Noguchi, M.; Tanaka, T.; Gyakushi, H.; Kobayashi, A.; Shoda, S. I., *J. Org. Chem.* **2009,** *74* (5), 2210-2212.

2. Tanaka, T.; Huang, W. C.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Tetrahedron Lett.* **2009,** *50* (18), 2154-2157.

3. Tanaka, T.; Nagai, H.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Commun.* **2009,** (23), 3378-3379.

4. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J., *Angew. Chem., Int. Ed.* **2014,** *53* (44), 1907-11911.

5. Tanaka, T.; Matsumoto, T.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Lett.* **2009,** *38* (5), 458-459.

6. Yoshida, N.; Noguchi, M.; Tanaka, T.; Matsumoto, T.; Aida, N.; Ishihara, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Asian. J.* **2011,** *6* (7), 1876-1885.

7. Yoshida, N.; Fujieda, T.; Kobayashi, A.; Ishihara, M.; Noguchi, M.; Shoda, S. I., *Chem. Lett.* **2013,** *42* (9), 1038-1039.

8. Novoa, A.; Barluenga, S.; Serba, C.; Winssinger, N., *Chem. Commun.* **2013,** *49* (69), 7608-7610.

9. Lee, Y. C.; Stowell, C. P.; Krantz, M. J., *Biochemistry* **1976,** *15* (18), 3956-3963.

10. Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G., *Angew. Chem., Int. Ed.* **2004,** *43* (7), 828-833.

11. Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G., *Angew. Chem., Int. Ed.* **2006,** *45* (24), 4007-4011.

12. Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G., *Org. Biomol. Chem.* **2003,** *1* (21), 3642-3644.

13. Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G., *J. Am. Chem. Soc.* **2008,** *130* (15), 5052-5053.

14. McSweeney, L.; Dénès, F.; Scanlan, E. M., *Eur. J. Org. Chem.* **2016,** *2016* (12), 2080-2095.

15. Gamblin, D. P. The Synthesis of Neoglycoconjugates. D.Phil. Thesis, Oxford University, 2005.

16. MacDougall, J. M.; Zhang, X.-D.; Polgar, W. E.; Khroyan, T. V.; Toll, L.; Cashman, J. R., *J. Med. Chem.* **2004,** *47* (23), 5809-5815.

17. Lim, D. University of Canterbury. Unpublished work, 2013.

18. Lim, D.; Fairbanks, A. J., *Chem. Sci.* **2017,** *8* (3), 1896-1900.

19. Buszewski, B.; Noga, S., *Anal. Bioanal. Chem.* **2012,** *402* (1), 231-247.

20. Behrend, R.; Roth, P., *Justus Liebigs Annalen der Chemie* **1904,** *331* (3), 359-382.

21. Zemplen, G.; Kunz, A., *Ber. Dtsch. Chem. Ges.* **1923,** *56*, 1705-1710.

22. Healy, J.; Ekkerman, S.; Pliotas, C.; Richard, M.; Bartlett, W.; Grayer, S. C.; Morris, G. M.; Miller, S.; Booth, I. R.; Conway, S. J., *Biochemistry* **2014,** *53* (12), 1982-1992.

23. Hoppmann, C.; Kühne, R.; Beyermann, M., *Beilstein J. Org. Chem.* **2012,** *8*, 884.

24. Comb, D. G.; Roseman, S., *J. Biol. Chem.* **1960,** *235* (9), 2529-2537.

25. Ma, P.; Schillinger, O.; Schwarten, M.; Lecher, J.; Hartmann, R.; Stoldt, M.; Mohrlüder, J.; Olubiyi, O.; Strodel, B.; Willbold, D.; Weiergräber, O. H., *Biochemistry* **2015,** *54* (35), 5469-5479.

26. Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G., *Angew. Chem., Int. Ed.* **2009,** *48* (42), 7798-7802.

27. Fischer, H.; Baer, R.; Hany, R.; Verhoolen, I.; Walbiner, M., *J. Chem. Soc. Perk. T. 2* **1990,** (5), 787-798.

28. Dondoni, A.; Massi, A.; Nanni, P.; Roda, A., *Chem. Eur. J.* **2009,** *15* (43), 11444-11449.

29. Ito, H.; Kamachi, T.; Yashima, E., *Chem. Commun.* **2012,** *48* (45), 5650-5652.

30. Cochrane, S. A.; Huang, Z.; Vederas, J. C., *Org. Biomol. Chem.* **2013,** *11* (4), 630-639.

31. Yamada, S.; Sudo, A.; Goto, M.; Endo, T., *RSC Advances* **2014,** *4* (56), 29890-29896.

32. Fiore, M.; Lo Conte, M.; Pacifico, S.; Marra, A.; Dondoni, A., *Tetrahedron Lett.* **2011,** *52* (3), 444-447.

33. Buchanan, C. M.; Phillips, A. R. J.; Cooper, G. J. S., *Biochem. J.* **2001,** *360* (2), 431-439.

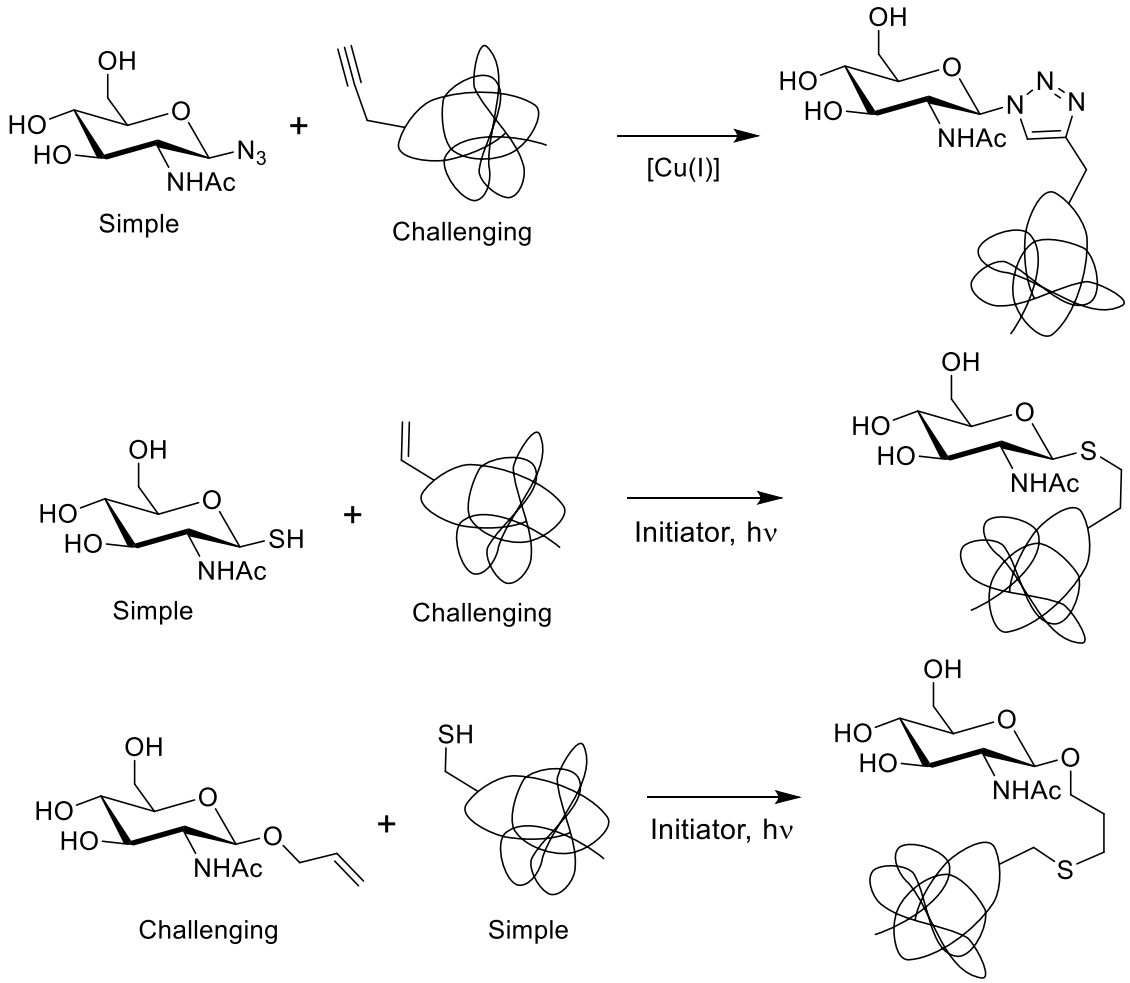
34. Seko, A.; Koketsu, M.; Nishizono, M.; Enoki, Y.; Ibrahim, H. R.; Juneja, L. R.; Kim, M.; Yamamoto, T., *Biochim. Biophys. Acta Gen. Subjects* **1997,** *1335* (1), 23-32.

35. Sun, B.; Bao, W.; Tian, X.; Li, M.; Liu, H.; Dong, J.; Huang, W., *Carbohydr. Res.* **2014,** *396*, 62-69

Chapter 4: Allyl propargyl ether as a bifunctional linker for the preparation of glycoconjugates

4.1 Limitations of current methods for glycoconjugate preparation

As discussed previously, most methods for the preparation of glycoproteins and glycopeptides suffer from one or more significant limitations, the most common of which is that synthesis of either the glycan or protein component presents a substantial challenge (**Scheme 4.1**).



Scheme 4.1: Methods of preparing glycoproteins.

Only a limited range of functional groups, represented by the standard proteogenic amino acids, can simply be introduced into a protein. Functional groups such as alkenes and alkynes are

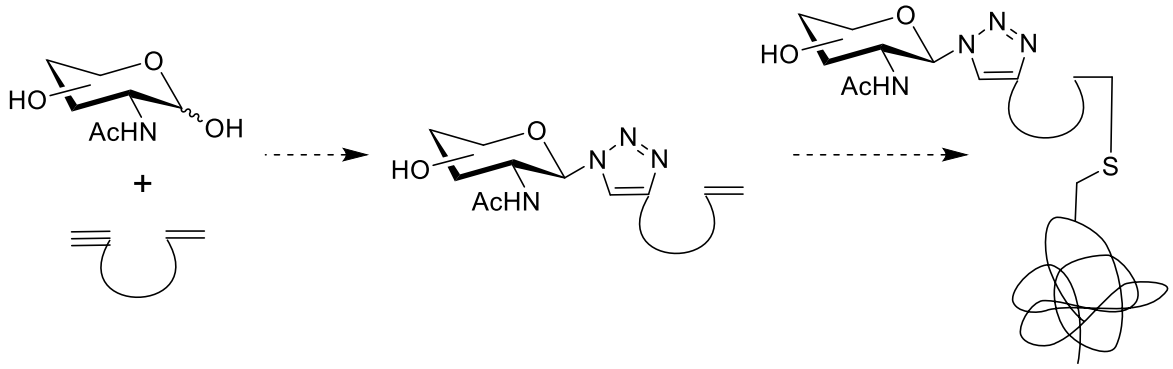
more difficult to introduce into a protein,[1] as they are not contained in the side chains of any of the amino acids that are coded for in standard protein biosynthesis. While an alkene or alkyne can be introduced into a protein it is non-trivial to do so.[2] In contrast, it is relatively easy to introduce reactive functional groups such as thiols through site-directed mutagenesis. Thiols are an effective handle for the site-selective modification of proteins, as they are both reactive and relatively uncommon, as most cysteine residues form disulfide linkages.

Selective chemical modification of large oligosaccharides also presents a significant challenge. As discussed previously (**Section 1.8**), DMC provides an effective means of modifying the reducing terminus of large carbohydrates without resorting to protecting groups. However, only a small variety of nucleophiles have been demonstrated to attack DMC-activated sugars, and introducing functional groups that can react with thiols, such as alkenes and alkynes, is at present non-trivial.[3, 4] While DMC-mediated reactions can be used to couple unprotected reducing sugars and thiols,[5, 6] the substrate scope is often limited, both in terms of thiol and sugar.

As no method currently exists in which preparation of both the sugar and protein component is simple, general site-selective preparation of glycoproteins remains a challenge.

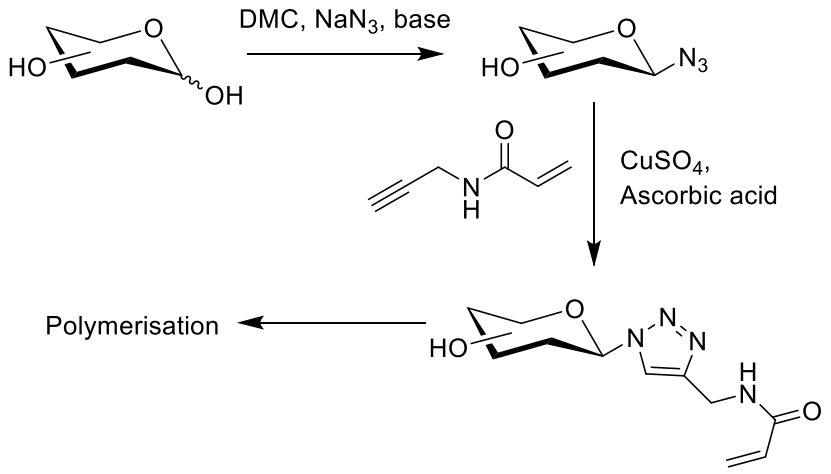
4.2 A bifunctional linker

Click reactions are perhaps the most useful reactions for the modification of proteins, being both robust and selective, but no single click reaction allows the simple preparation of glycoproteins without challenging alteration of either the sugar or the protein component. However, it is possible to imagine that a combination of click reactions could be effective for the preparation of glycoproteins. For example, CuAAC and thiol-ene click chemistry occur under different conditions, and, most significantly, alkenes are entirely unreactive in CuAAC. However, the reverse is not true; alkynes can react in thiol-ene click reactions.[7] Therefore, if CuAAC was used to attach an alkene to a carbohydrate, it would then be possible to prepare a glycoprotein or glycopeptide using a thiol-ene click reaction (**Scheme 4.2**) in a second step. This process would require minimal manipulation of the carbohydrate, and only the introduction of a cysteine residue into the protein or peptide.



Scheme 4.2: The use of a bifunctional linker to prepare a glycoprotein.

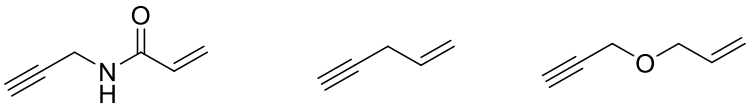
The CuAAC reaction can be easily carried out on complex reducing sugars in one pot using the method previously developed by Lim in the Fairbanks group,[8] while the thiol-ene click reaction has previously been demonstrated to be effective for the preparation a variety of glycoconjugates.[9]



Scheme 4.3: Attachment of an alkene to a carbohydrate by Shoda *via* CuAAC.[4]

The use of CuAAC to introduce an alkene is similar to the method developed by Shoda to create glycopolymers (**Scheme 4.3**),[4] but the use of a reagent with an alkyne at one end and an alkene at the other to prepare glycoconjugates through the thiol-ene click reaction has not been attempted.

4.3 Choice of linker

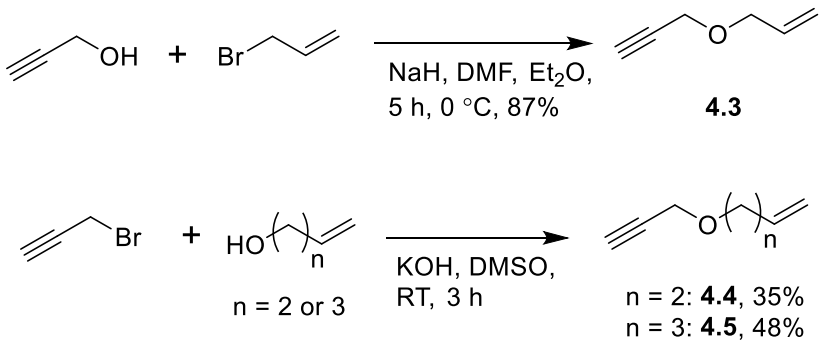


4.1 **4.2** **4.3**

Figure 4.1: Potential bifunctional linkers.

In his previous study, Shoda used propargyl acrylamide **4.1** (**Figure 4.1**) as a linker molecule which contains both an alkene and an alkyne.[4] While this does contain both an alkene and an alkyne, the use of an acrylamide linker was presumably due to the subsequent use of the sugar-linker conjugates for polymerisation reactions, where acrylamide monomers are widespread. Linker **4.1** is commercially available, but not from major suppliers. It can be prepared, but either expensive reagents or multiple steps are required.[10, 11]

A number of other linkers were therefore considered. The ideal linker would be simple, readily available, and easily modifiable. An alkyl linker such as **4.2** is simple, but presents a significant challenge to synthesise. An ether linker such as **4.3** is simple to prepare using standard techniques in a manner that allows variation,[12, 13] and is stable under biological conditions. Furthermore, linker **4.3** is readily commercially available.

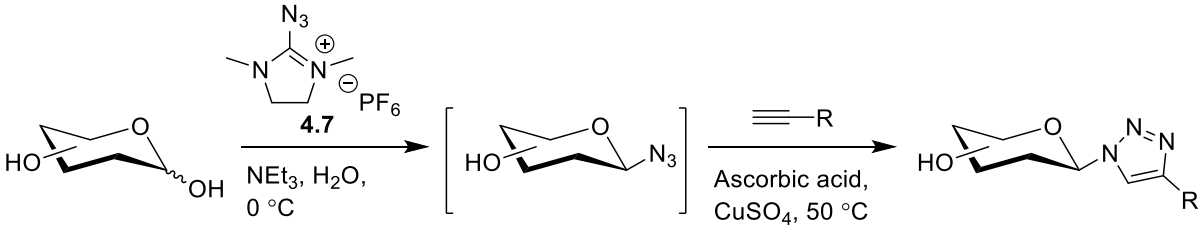


Scheme 4.4: Preparation of ether linkers.

With a view to testing the effectiveness of linkers of various lengths in CuAAC and the subsequent thiol-ene reaction, three propargyl ether based-linkers **4.3**-**4.5** were prepared, which differed in the length of the alkyl chain between the alkene and oxygen. All were prepared using a broadly similar procedure in which an alkyl bromide was reacted with an alcohol in the presence of a base. Allyl propargyl ether **4.3** was prepared following a literature procedure,[12] in which propargyl alcohol was reacted with allyl bromide using sodium hydride as the base (**Scheme 4.4**). An attempt to prepare butenyl propargyl ether **4.4** *via* the same method proved unsuccessful, as only a small amount of product was formed. An alternative literature procedure, in which potassium hydroxide was used as the base, propargyl bromide was used as the electrophile, and 3-buten-1-ol was used as the nucleophile, was successful.[13] This procedure afforded butenyl linker **4.4** in 35% yield. An analogous procedure using 4-penten-1-ol as the nucleophile afforded pentenyl linker **4.5** in a yield of 48%. Both ethers are volatile,

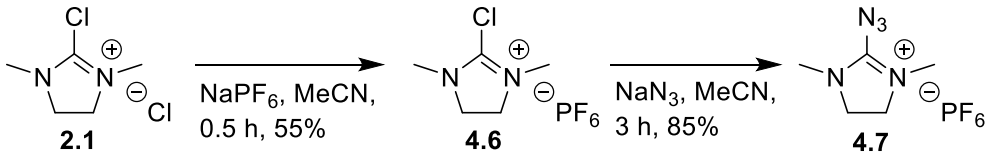
and the low yield of product can be explained by the difficulty of completely separating solvent from the product ether.

4.4 One pot preparation of sugar-linker conjugates



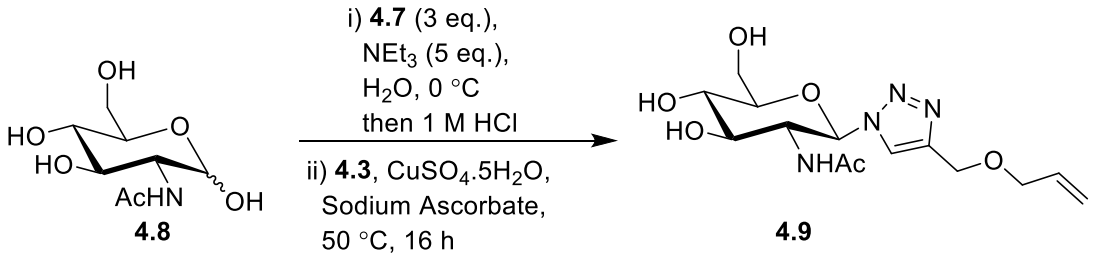
Scheme 4.5: One pot conversion of reducing sugars to glycosyl azides followed by CuAAC.

With a selection of linkers available the next step was to conjugate a simple sugar to a linker. As discussed previously (**Section 1.7.2**), Lim developed the one-pot conversion of unprotected reducing sugars to glycosyl azides and their use in CuAAC in the Fairbanks group (**Scheme 4.5**),[8] and this methodology was ideal for applying to the linkers developed here.



Scheme 4.6: Preparation of ADMP.

With the aim of applying this reaction to linkers **4.3-4.5**, ADMP **4.7** was prepared from DMC **2.1** following the literature procedure (**Scheme 4.6**).[8] First, DMC **2.1** was stirred with sodium hexafluorophosphate to afford hexafluorophosphate salt **4.6** in 55% yield, which was then reacted with sodium azide to produce ADMP **4.7** in 85% yield.



Scheme 4.7: ADMP-mediated reaction of GlcNAc and linker **4.3**.

Solvent for Part ii)	$CuSO_4$ Equivalents	Sodium Ascorbate eq.	Equivalents of ether 4.3	% Conversion to 4.9
H_2O	0.015	0.2	1	trace
H_2O	0.1	0.2	1	18%
$^tBuOH:H_2O$ 2:1	0.1	0.2	1	68%
$^tBuOH:H_2O$ 2:1	0.1	0.2	2	72%
$^tBuOH:H_2O$ 2:1	0.2	0.4	2	86%

Table 4.1: Optimisation of the CuAAC click reaction.

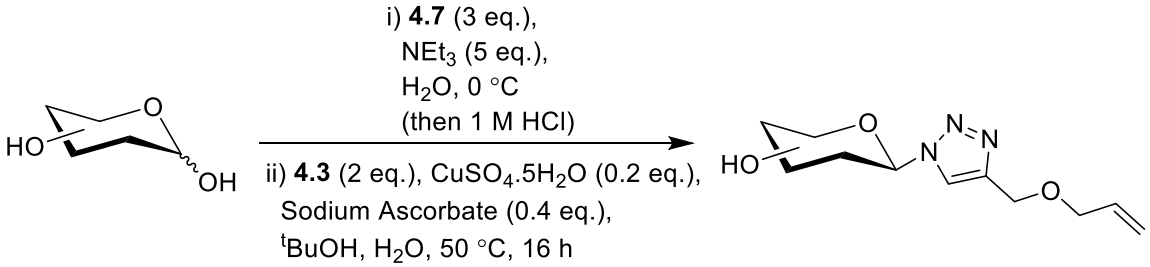
With all the required starting materials in hand, the reaction of GlcNAc **4.8** and linker **4.3** to produce glycosyl triazole **4.9** was examined (**Scheme 4.7**). As a starting point, the conditions of Lim were applied (**Table 4.1**), namely 3 equivalents of ADMP **4.7** and 5 equivalents of triethylamine in water at 0 °C, followed by the addition of 1 M hydrochloric acid, stirring at pH 2 for 10 min, neutralisation, then the addition of 1 equivalent of alkyne **4.3**, 0.015 equivalents of copper sulfate, and 0.2 equivalents of sodium ascorbate and heating in water at 50 °C.[8] Only a trace amount of product was formed under these conditions. Increasing the number of equivalents of copper sulfate resulted in the formation of a measurable amount of product, with a conversion to product **4.9** of 18% (as measured by ^1H NMR spectroscopy). Reasoning that perhaps the low conversion to product was due to the insolubility of the non-

polar ether **4.3** in water, a water/*tert*-butanol mixture was tested as a solvent, a change which resulted in a significantly higher conversion to **4.9** of 68%. Increasing the number of equivalents of all reagents again resulted in an increase in conversion to product, with the highest conversion to product of 86% observed when 0.2 equivalents of copper sulfate, 0.4 equivalents of sodium ascorbate, and 2 equivalents of alkyne **4.3** were used.

Using these optimised conditions, an isolated yield of GlcNAc-linker conjugate **4.9** of 79% was obtained for this reaction. Purification was straightforward using ion exchange filtration followed by flash column chromatography. However it was observed that an impurity which does not produce any signals in either ^{1}H or ^{13}C NMR spectra was a common contaminant. Careful purification was required to ensure that this contaminant was indeed removed. The observation of peaks in ^{19}F and ^{31}P NMR spectra of some samples of **4.9** corresponding to the hexafluorophosphate anion suggested that the impurity was sodium hexafluorophosphate, and ^{19}F and ^{31}P NMR spectroscopy were used to confirm that all sodium hexafluorophosphate was removed from purified samples of **4.9**. The use of an internal standard when taking ^{1}H NMR spectra of **4.9** further confirmed that all non-NMR active impurities had been removed.

4.4.1 Variation of the sugar substrate

Having demonstrated that the CuAAC click reaction component of the sequential double click reaction approach could be successfully carried out, a variety of carbohydrates were then attached to linker **4.3** (**Scheme 4.8, Table 4.2**).



Scheme 4.8: Attachment of sugars to alkyne **4.3**.

Pleasingly, in most cases the conditions developed for the attachment of GlcNAc to linker **4.3** worked equally well for other sugars. Treatment with 1 M hydrochloric acid was only required for 2-acetamido sugars in order to open the 1,2-oxazoline that formed competitively upon ADMP activation of the sugar, and was not necessary for 2-hydroxy sugars.

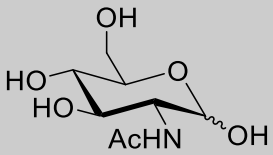
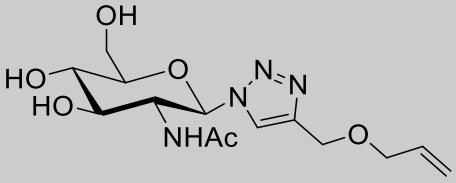
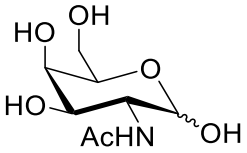
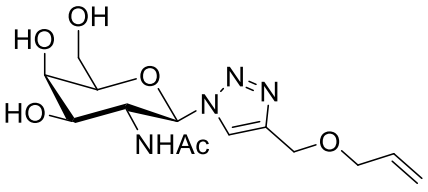
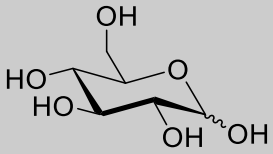
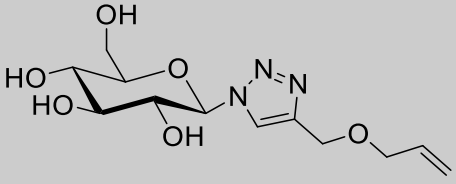
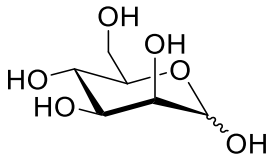
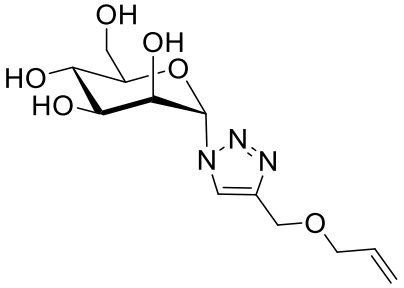
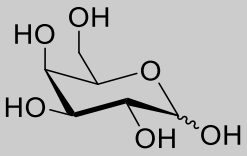
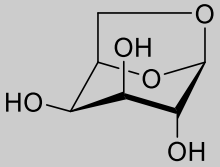
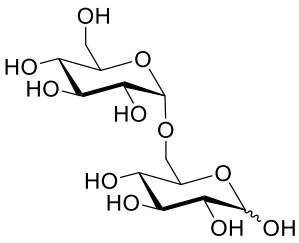
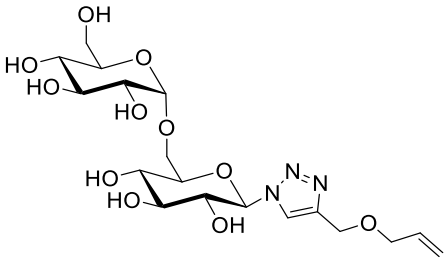
Sugar	Product	Isolated Yield
	 **4.9**	79%[a]
	 **4.10**	44%[a]
	 **4.11**	66%
	 **4.12**	80%[b]
	 **4.13**	41% (85% conversion to 1,6-anhydrogalactose)
	 **4.14**	75%

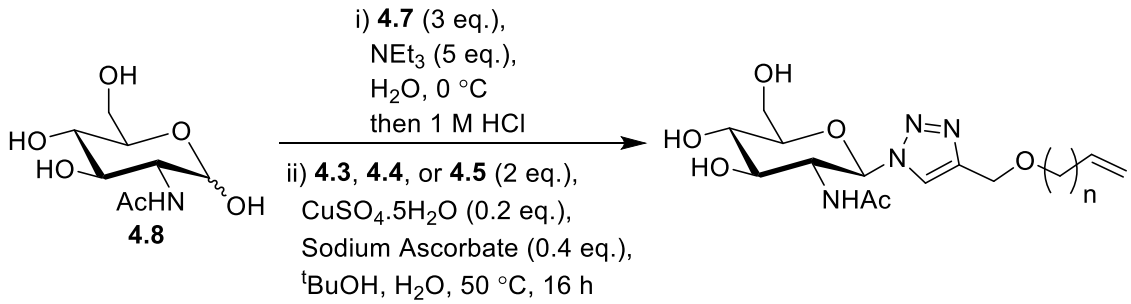
Table 4.2: Attachment of various sugars to linker **4.3**. [a]Addition of 1 M HCl was required to open oxazoline intermediate. [b]Reaction time: 40 h.

The use of GalNAc as a substrate resulted in a reduced yield of product **4.10** of 44%, but as discussed previously (**Section 2.4**) this observation is consistent with previous DMC- and

ADMP-mediated reactions where GalNAc or galactose has been used as a substrate. Glucose and mannose both produced the desired click product in moderate to high yields, (66% and 80% respectively). The click reaction of mannose required an extended reaction time. Wilkinson and co-workers previously observed that α-glycosyl azides react more slowly than β-glycosyl azides in CuAAC click reactions.[14] This observation was later examined in more detail by Dedola *et al*.,[15] who concluded that the difference in reactivity between the two was significant, as α-glycosyl azides reportedly react more than three times more slowly in CuAAC than β-azides. The difference in reaction rates has been proposed to be due to the anomeric effect, which results in a decrease in the molecular orbital coefficient in the HOMO of the p-orbital of the anomeric nitrogen of the azide functional group. Practically, this altering of the azide molecular orbitals lowers the affinity of the anomeric nitrogen for copper(I) ions, resulting in a reduced reaction rate.[16]

Unfortunately, galactose did not prove compatible with the ADMP-mediated reaction. Rapid formation of 1,6-anhydrogalactose **4.13** was observed instead of the formation of the desired galactosyl azide. 1,6-Anhydrogalactose **4.13** was the major product, and formed in a ratio of approximately 6:1 with the desired click product. The use of galactose in a reaction mediated by either DMC or ADMP to produce a galactosyl azide has never been reported,[8, 17] presumably due to the unavoidable formation of 1,6-anhydrogalactose **4.13**. As previous investigations using similar methodology apparently did not succeed in this area, galactose was abandoned as a substrate for this reaction. More promisingly, reaction of isomaltose with linker **4.3** produced the desired product **4.14** in 75% yield, proving that the reaction is effective with sugars larger than a monosaccharide.

4.4.2 Variation of the bifunctional linker



Scheme 4.9: Attachment of linkers of a range of lengths to GlcNAc **4.8**.

Variation of the linker used would allow fine-tuning of the properties of any glycoconjugates prepared by the sequence of click reactions described in this chapter. In order to demonstrate that the linker could be varied without drastically altering the nature of the click reaction, the extended linkers **4.4** and **4.5** were examined (**Scheme 4.9**).

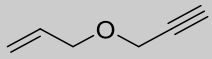
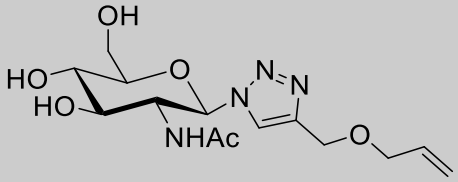
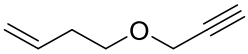
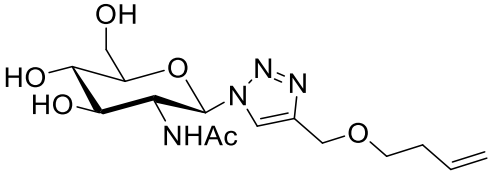
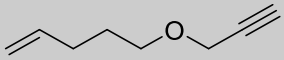
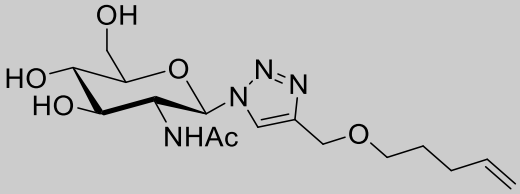
Linker	Product	Isolated Yield
 **4.3**	 **4.9**	79%
 **4.4**	 **4.15**	53%
 **4.5**	 **4.16**	89%

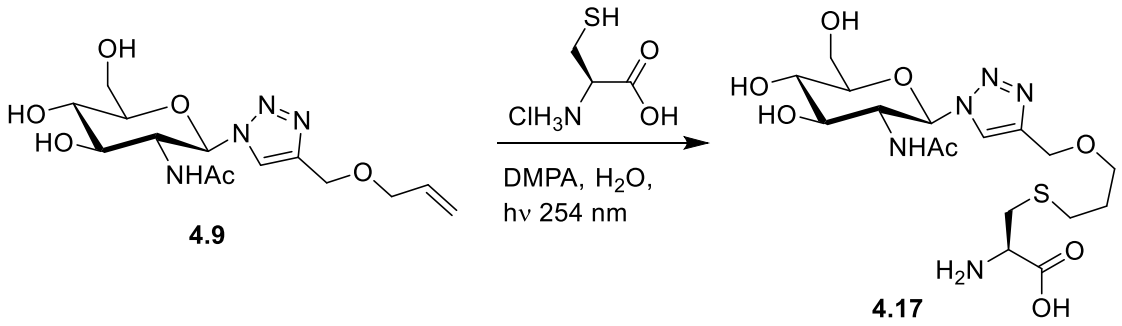
Table 4.3: Attachment of linkers of a range of lengths to GlcNAc **4.8**.

Pleasingly, variation of the linker length resulted in uniformly high conversions in all cases (**Table 4.3**). The lowest yield was observed for the reaction of butenyl linker **4.4**. However the

conversion to product as assessed by ^{1}H NMR spectroscopy of the crude reaction mixture was >90% in all cases, implying that the issue in this case was purification of the product **4.15** rather than incomplete reaction.

4.5 Application of the thiol-ene click reaction to the glycosyl triazoles

With conditions established for the attachment of carbohydrates to one end of the bifunctional linker, the next step was to confirm that it is then possible to attach a thiol to the other end of the linker. Cysteine was selected for initial studies, as the primary application envisaged for this reaction was the conjugation of sugars to the cysteine residues of peptides or proteins.



Scheme 4.10: Thiol-ene click reaction of alkene **4.9** with cysteine.

The initial conditions selected were those that previously proved effective for the thiol-ene click reaction of glycosyl thiols (**Section 3.8.2**), i.e. 1 equivalent of cysteine, 0.1 equivalents of DMPA, water as the solvent, irradiation at 254 nm, and a reaction time of 6 h (**Scheme 4.10**). Disappointingly, a low conversion to product **4.17** was observed for this substrate combination (**Table 4.4**).

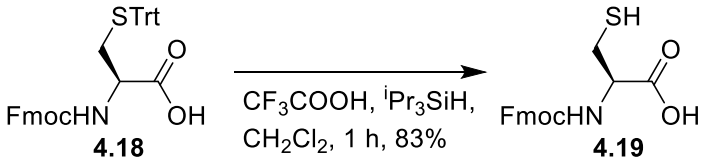
Reaction Time	Cysteine Equivalents	DMPA Equivalents	Under N_2?	Solvent	Conversion
6 h	1	0.1	No	H_2O	25%
16 h	1	0.1	No	H_2O	37%
6 h	0.5	0.1	No	H_2O	30%[a]
6 h	2	0.1	No	H_2O	37%
6 h	1	1	No	H_2O	32%
6 h	1	0.1	Yes	H_2O	42%
6 h	1	0.1	No	pH 4.0 Buffer[b]	40%

Table 4.4: Testing of conditions for the thiol-ene click reaction of **4.9** with cysteine. [a]Conversion calculated with respect to cysteine. [b]pH 4.0 0.2 M Sodium acetate buffer

Increasing the number of equivalents of either cysteine or alkene, or increasing the reaction time, resulted in only a slight improvement. Increasing the number of equivalents of the radical initiator DMPA used did not significantly improve the conversion, even when a stoichiometric amount of initiator was used. Carrying out the reaction under N_2 resulted in a slight improvement, but the reaction remained low-yielding. The use of a pH 4.0 sodium acetate buffer as a solvent has previously been shown to increase the efficiency of the thiol-ene click reaction,[18] and while in this case a slight improvement was observed, the conversion to product remained disappointingly low. Testing of the reaction mixture with Ellman's reagent confirmed that while some oxidation to cystine occurred over the course of the reaction (approximately 20% of the added cysteine was converted to cystine), the major issue was simply the low reactivity of cysteine with alkene **4.8**, as a significant portion of unreacted cysteine remained.

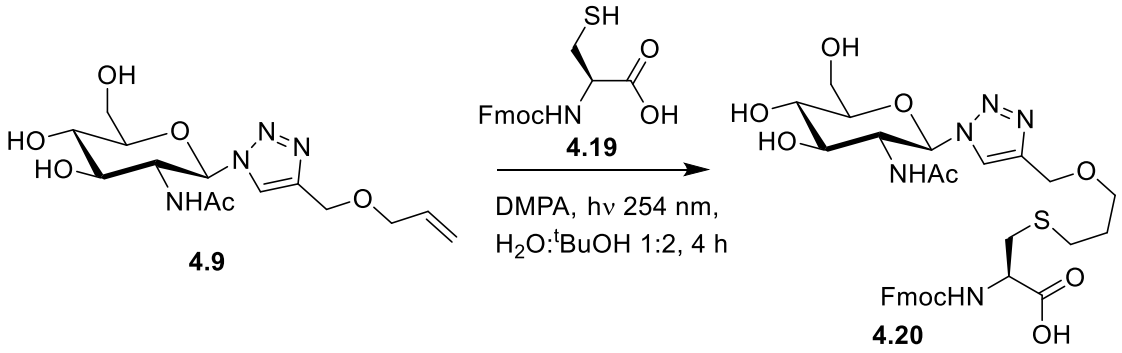
With the low reactivity of cysteine in mind, the next step was to investigate the reactivity of other, similar, thiols. Fmoc-Protected cysteine **4.19** was prepared by deprotection of

commercially available Fmoc-*S*-tritylcysteine **4.18** using trifluoroacetic acid and triisopropylsilane (**Scheme 4.11**).



Scheme 4.11: Preparation of Fmoc-protected cysteine.

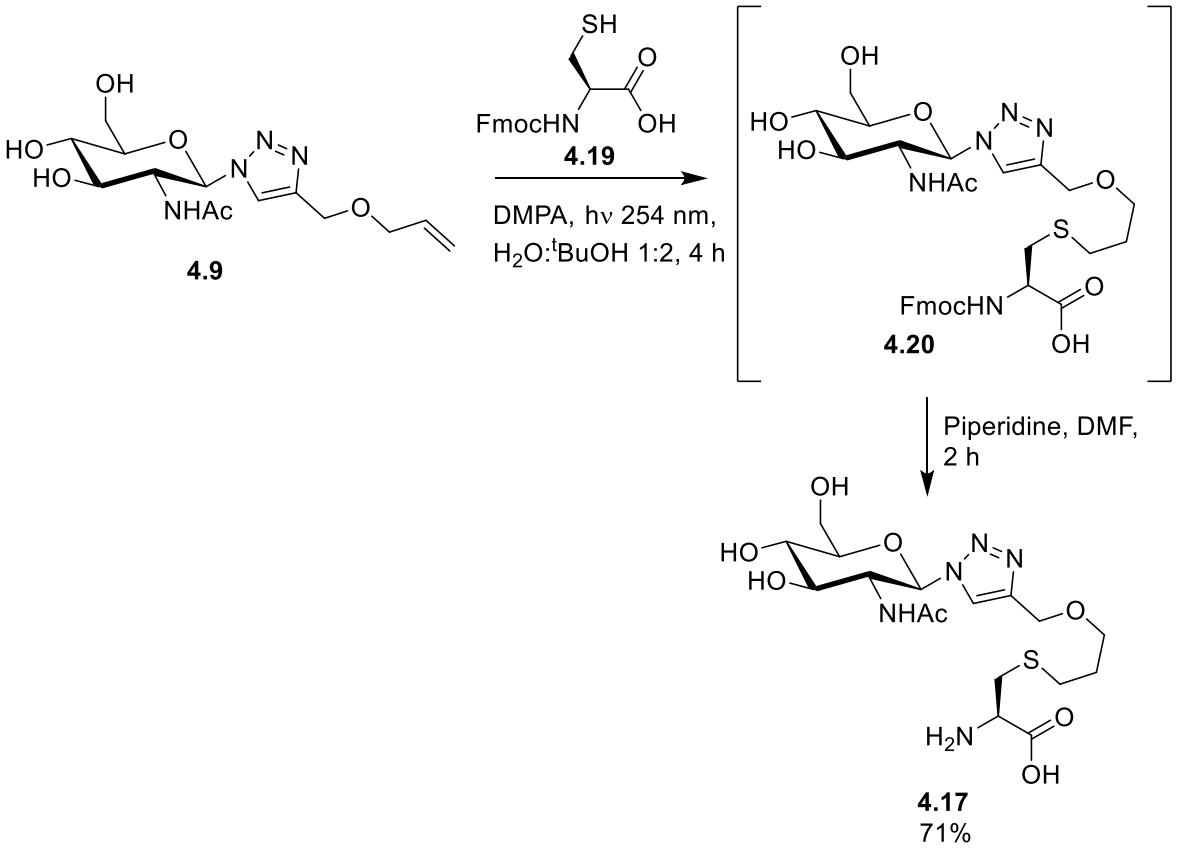
An initial investigation demonstrated that Fmoc-cysteine **4.19** was significantly more reactive than cysteine in the thiol-ene click reaction (**Scheme 4.12**). 0.1 Equivalents of DMPA and 1 equivalent of Fmoc-cysteine **4.19** were used, and the reaction was irradiated at 254 nm for 6 h. A *tert*-butanol/water mixture (2:1) was used as the solvent because Fmoc-cysteine **4.19** is insoluble in water alone.



Scheme 4.12: Thiol-ene click reaction using Fmoc-cysteine.

However, the reaction proved highly inconsistent, with the conversion to product varying between 50% and >95% for no discernible reason. An investigation into the oxidation behaviour of Fmoc-cysteine **4.19** under the reaction conditions revealed that Fmoc-cysteine **4.19** was completely oxidised to disulfide under the reaction conditions, in a process that was significantly accelerated by radical initiator DMPA and irradiation with UV light. Previous studies have shown that the free-radical oxidation of thiols is accelerated by the presence of

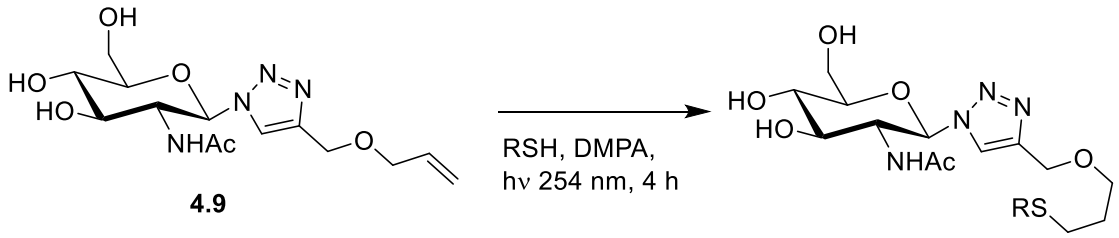
oxygen,[19] and therefore a test oxidation reaction under an atmosphere of nitrogen was carried out, which resulted in reduced oxidation of Fmoc-cysteine. This observation implied that if the thiol-ene click reaction was carried out under an atmosphere of nitrogen the thiol-ene click reaction would be more effective and consistent, and this did indeed prove to be the case. The use of two equivalents of Fmoc-cysteine **4.19** resulted in complete conversion to glyco-amino acid **4.20**.



Scheme 4.13: Purification of glycoconjugate **4.17** after Fmoc deprotection.

Purification of glyco-amino acid **4.20** was challenging due to its amphiphilic nature, so in order to simplify the process the Fmoc protecting group was removed and the product **4.17** was purified by reverse phase flash column chromatography (**Scheme 4.13**). An overall yield of 71% was obtained for the transformation from alkene **4.9**.

The above experiments revealed two potential issues with the thiol-ene click reaction of alkene **4.9** and amino acid thiols: either the thiol did not react at a high enough rate with the alkene, or the thiol was easily oxidised to disulfide. While no means of overcoming the first issue has been discovered, the second was relatively easily avoided by conducting the reaction under an atmosphere of nitrogen and using excess thiol. With this in mind, a small selection of other thiols were investigated for their utility in this thiol-ene click reaction (**Scheme 4.14**). The use of a pH 4.0 0.2 M sodium acetate buffer resulted in an increase in conversion to product in some cases.



Scheme 4.14: Thiol variation in the thiol-ene click reaction of alkene **4.9**.

In all cases a moderate to high yield was obtained (**Table 4.5**). The variation in the number of equivalents of thiol used reflected the ease of oxidation of the thiol starting material; if the thiol was more readily oxidised a greater number of equivalents of thiol were required to ensure that the majority of the sugar alkene reacted.

The highest conversion and yield of isolated product was obtained when glutathione **4.17** was used as the thiol, a promising result because glutathione was closer to the intended thiol targets of this research, namely peptides and proteins. The fact that the tripeptide glutathione showed significantly greater activity in the thiol-ene click reaction with alkene **4.9** as compared to single amino acids suggested that the bifunctional click reagent approach should also be useful for the preparation of glycopeptides and glycoproteins.

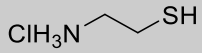
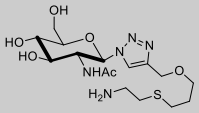
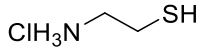
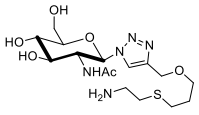
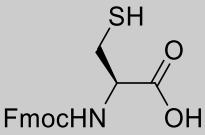
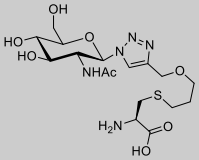
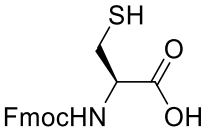
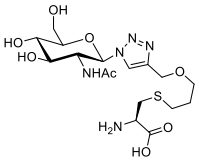
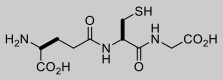
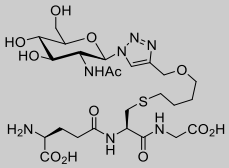
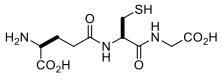
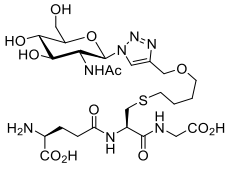
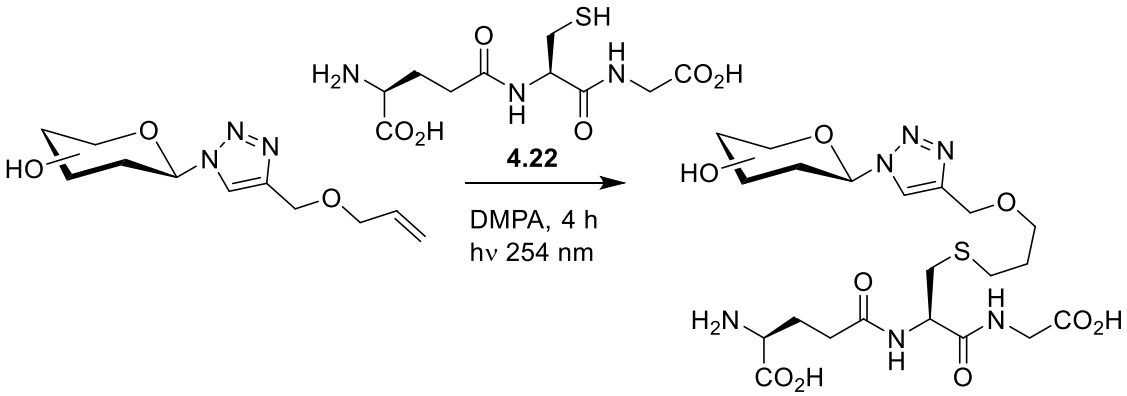
Thiol	Product	Solvent	Thiol Equivalents	Conversion	Yield
	 **4.21**	H_2O	3	78%	-
	 **4.21**	pH 4.0 sodium acetate buffer	3	90%	61%
 **4.19**	 **4.17**	$^tBuOH:H_2O$ 2:1	2	quant.	71%
 **4.19**	 **4.17**	$^tBuOH:H_2O$ 2:1, pH 4.0 sodium acetate buffer	2	quant.	74%
 **4.22**	 **4.23**	H_2O	1	58%	-
 **4.22**	 **4.23**	pH 4.0 sodium acetate buffer	1	quant.	quant.

Table 4.5: Variation of thiol in the thiol-ene click reaction of alkene **4.9**.

4.5.1 Variation of the sugar alkene in the thiol-ene click reaction

Having demonstrated that the thiol-ene click reaction could be applied to sugar alkenes prepared *via* CuAAC, the next step was to confirm that the thiol-ene click reaction was effective using the alkene-containing sugars prepared in this manner. Glutathione **4.22** was selected as a model thiol substrate, as it was both a highly efficient in the thiol-ene click reaction and representative of larger peptide thiols, which are more interesting targets from a biological perspective.



Scheme 4.15: Variation of the sugar-alkene substrate in the thiol-ene click reaction.

The thiol-ene click reaction with glutathione **4.17** (**Scheme 4.15**) proved effective with all of the sugar alkenes that had been prepared (**Table 4.6**). The isolated yields of the glycopeptides produced varied from moderate to very high.

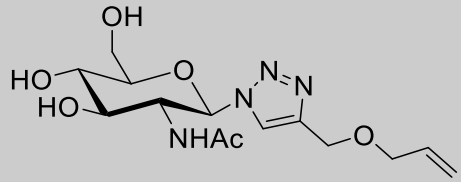
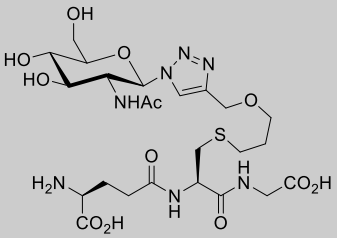
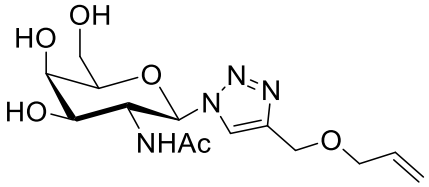
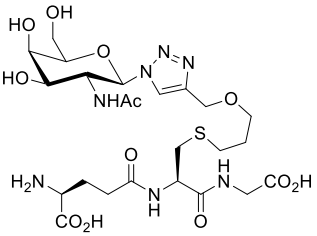
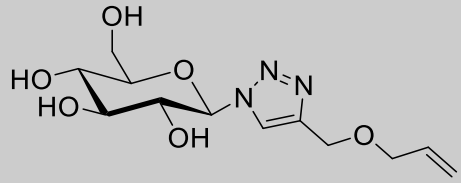
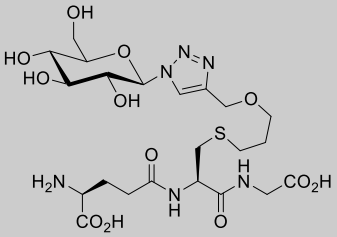
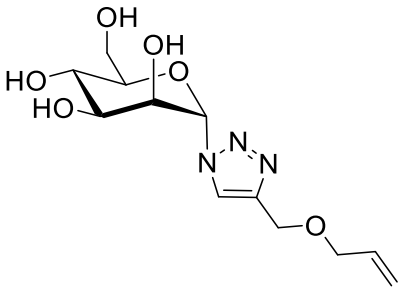
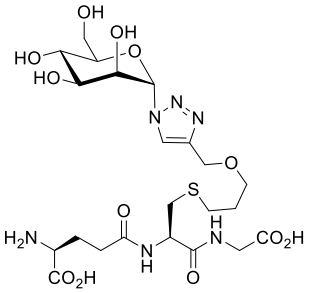
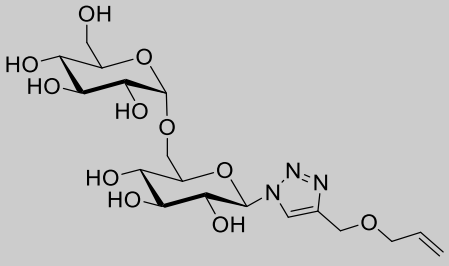
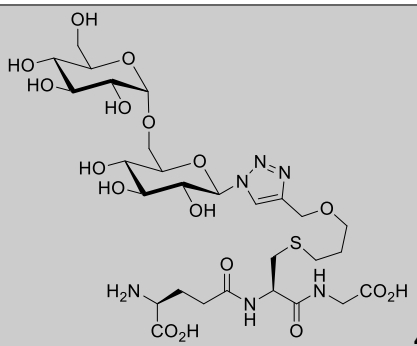
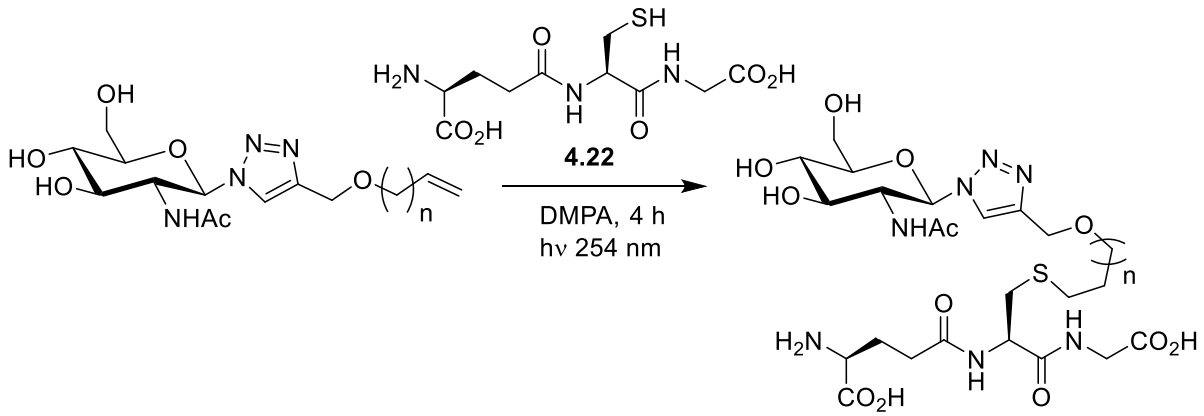
Sugar	Product	Isolated Yield
 **4.9**	 **4.23**	quant.
 **4.10**	 **4.24**	70%
 **4.11**	 **4.25**	93%
 **4.12**	 **4.26**	74%
 **4.14**	 **4.27**	76%

Table 4.6: Thiol-ene click reaction of various sugar-alkenes with glutathione **4.17**.

4.5.2 Variation of the linker length in the thiol-ene click reaction



Scheme 4.16: Variation of the linker length in the thiol-ene click reaction.

As expected, the thiol-ene click reaction proved effective when an extended linker was used (**Scheme 4.16, Table 4.7**); butyl-linked glycopeptide **4.28** and pentyl-linked glycopeptide **4.29** were both obtained in moderate yield. The reduced yield in comparison to propyl-linked glycopeptide **4.16** was most likely due to the increased hydrophobicity of the linker resulting in greater losses during purification, as the conversion to product as determined by ^{1}H NMR of the crude reaction mixture was very high in all cases.

As variation of the sugar and thiol components of the reaction, as well as the linker length, had been tolerated in the reaction sequence developed, it appeared that the use of a bifunctional click linker represented a versatile technique for the conjugation of unprotected reducing sugars to thiols.

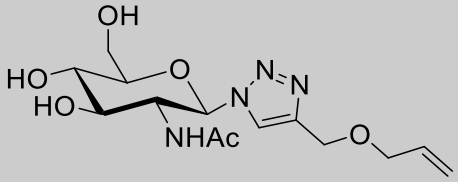
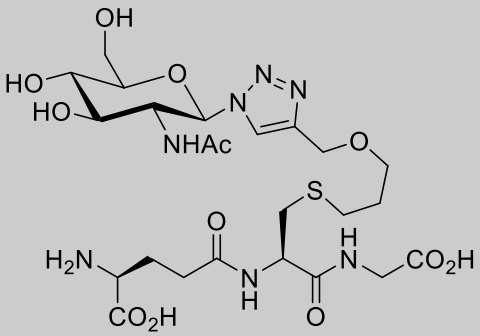
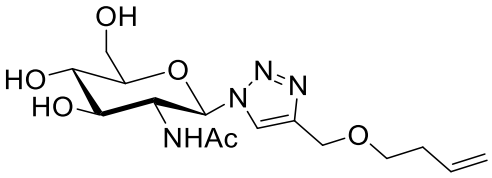
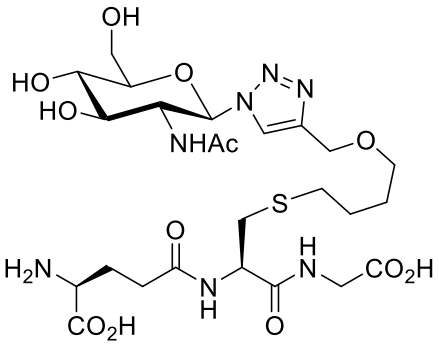
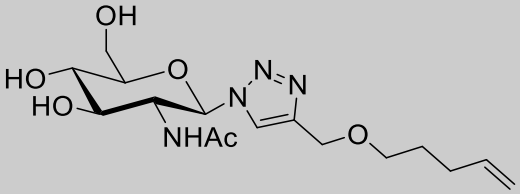
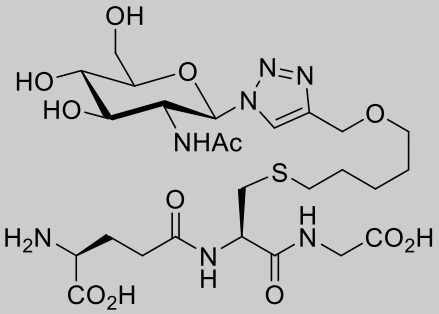
Starting material	Product	Isolated Yield
 **4.9**	 **4.18**	quant.
 **4.10**	 **4.28**	63%
 **4.11**	 **4.29**	75%

Table 4.7: Variation of the linker length in the thiol-ene click reaction.

4.6 Preparation of MUC1 glycopeptides

The use of a bifunctional linker has been shown to be a versatile method for the preparation of small glycopeptides. However, preparation of a more biologically relevant example was necessary in order to demonstrate its true utility.

As discussed in **Section 1.2.2** and **Section 1.6.1**, mucins are a class of glycoproteins commonly found as a component of mucous secretions and as transmembrane proteins.[20] Abnormal glycosylation of mucins is linked to some cancers,[21] and the preparation of mucin glycoconjugates is therefore of great interest for the design of anti-cancer vaccines.[22] Mucins contain a variable number tandem repeat (VNTR) domain composed of a repeating peptide unit which is heavily *O*-glycosylated, and short glycosylated peptides therefore represent prime vaccine targets.[22] The repeated peptide unit of MUC1, VTSAPDTRPAPGSTAPPAHG (**Figure 4.2**),[23] has five potential *O*-glycosylation sites, highlighted in red. Glycosylation of one or more of these sites increases the immunogenic activity of the peptide,[22] and MUC1 was therefore thought to be an ideal target for modification by the double click reaction methodology developed in this chapter.

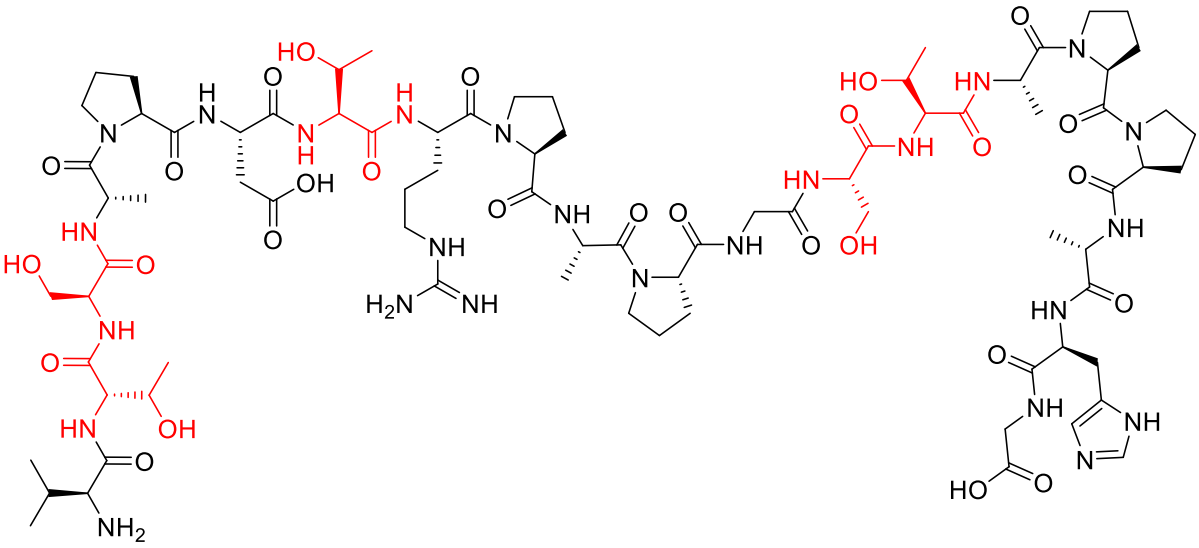


Figure 4.2: MUC1 VNTR domain repeat peptide unit, with *O*-glycosylation sites highlighted in red.

In order to modify the MUC1 repeat peptide with the double click method a cysteine needed to be introduced at one of the *O*-glycosylation sites. The threonine residue at position 7 was selected for replacement (**Figure 4.3**), as studies have shown that glycosylation at this position is significant for the immunogenic activity of the peptide.[22] Peptide **4.30** was therefore prepared by Dr Geoff Williams of the Brimble group at the University of Auckland using standard solid phase peptide synthesis.

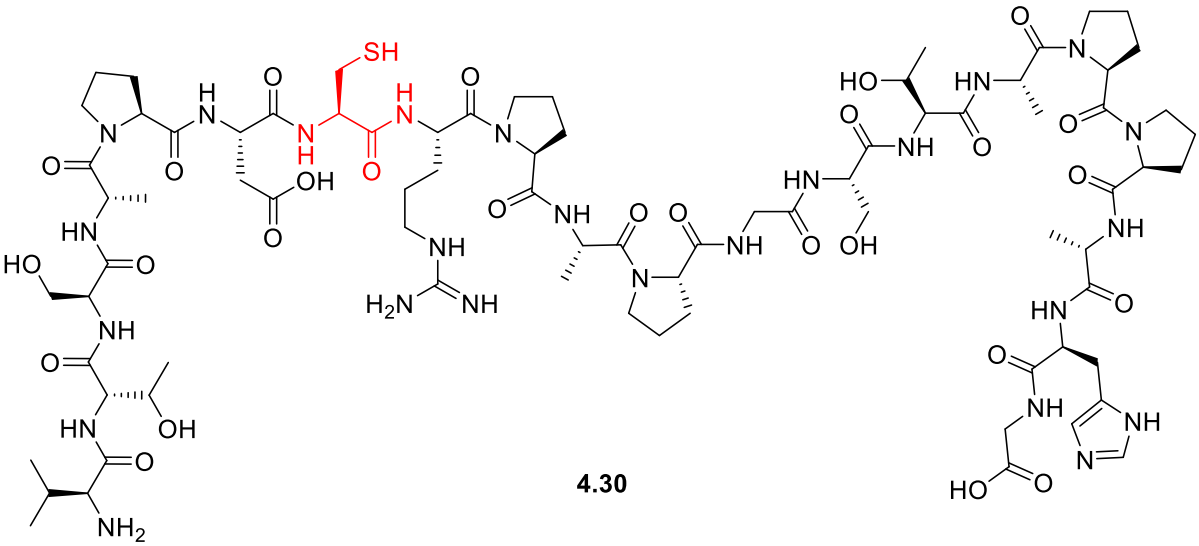
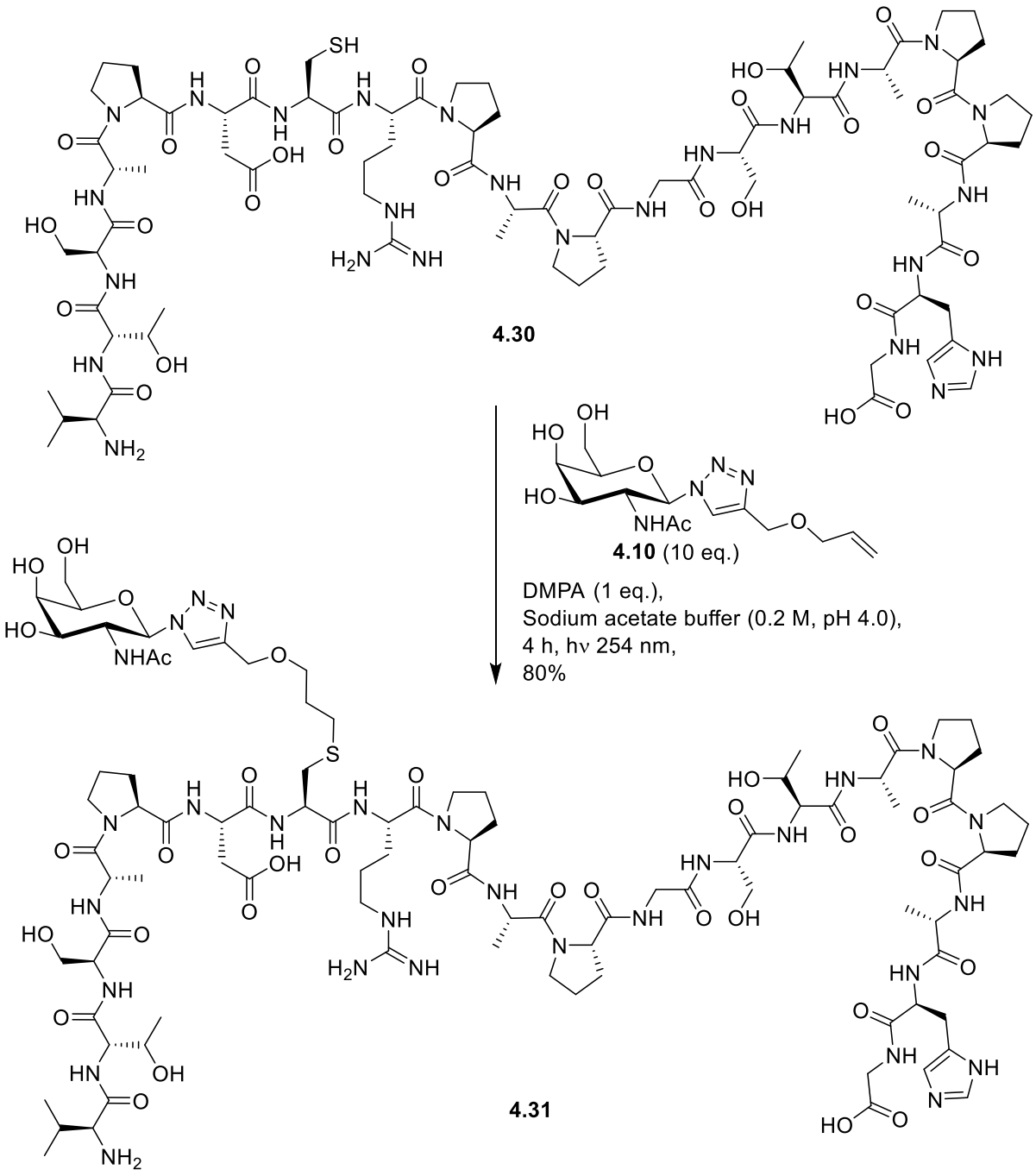


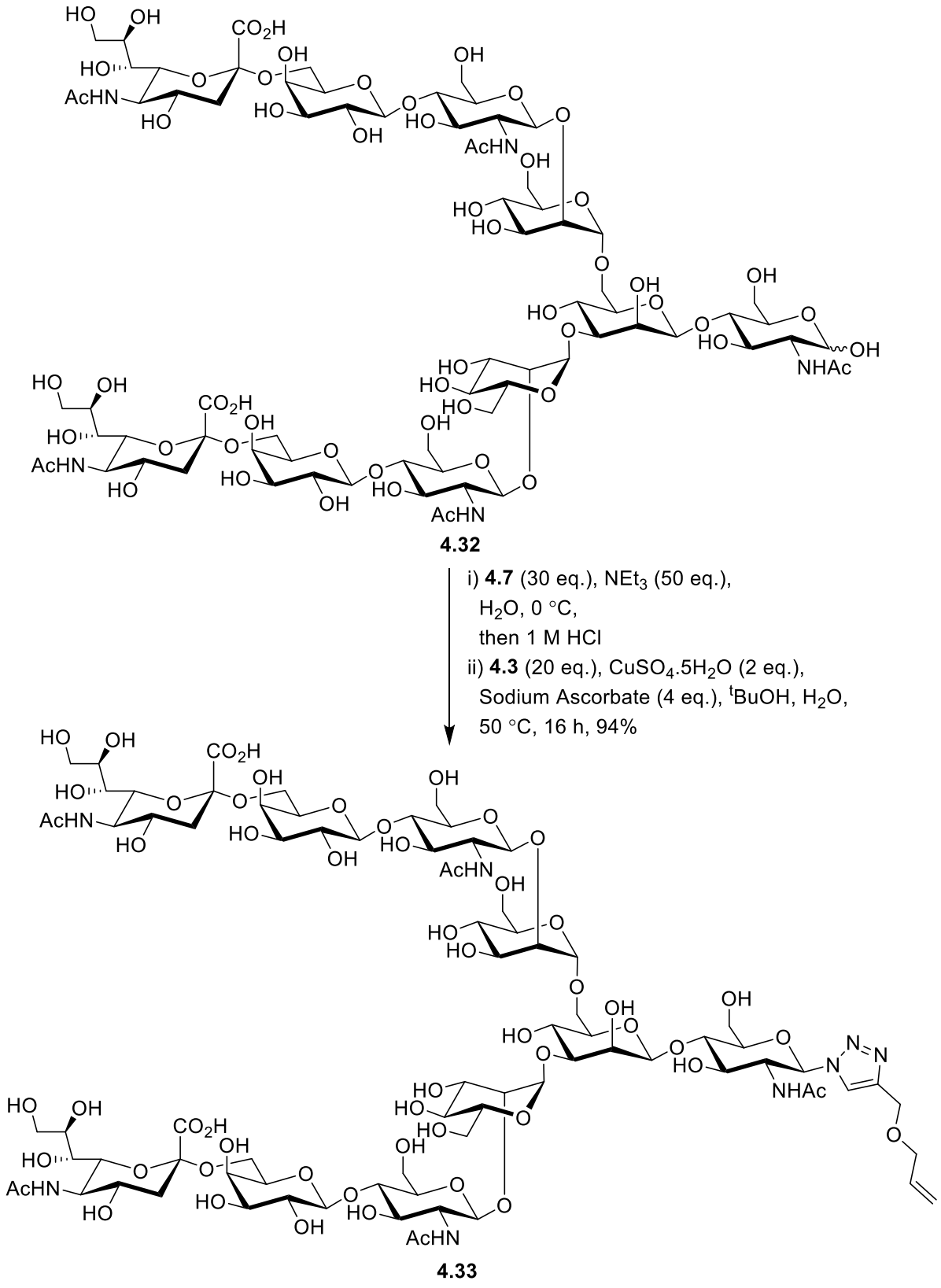
Figure 4.3: MUC1 20mer peptide with a cysteine inserted at position 7, highlighted in red.

The Tn antigen (α-GalNAc) is overexpressed on mucins in a variety of cancers,[24] therefore the GalNAc derivative **4.17** represents a suitable sugar for attachment to peptide **4.30** in order to prepare a biologically relevant glycopeptide.

Pleasingly, the thiol-ene click reaction of peptide thiol **4.30** and GalNAc alkene **4.17** (**Scheme 4.17**) proved effective, producing glycopeptide **4.31** in 80% yield, as determined by HPLC. An excess of GalNAc alkene **4.19** was required to ensure significant glycosylation of the peptide. The use of a 20mer peptide in the thiol-ene click reaction demonstrated that the use of two sequential click reactions on a bifunctional linker was an effective method for the preparation of biologically relevant glycoconjugates.



Scheme 4.17: Attachment of GalNAc triazole **4.10** to modified MUC1 20mer peptide **4.30**.

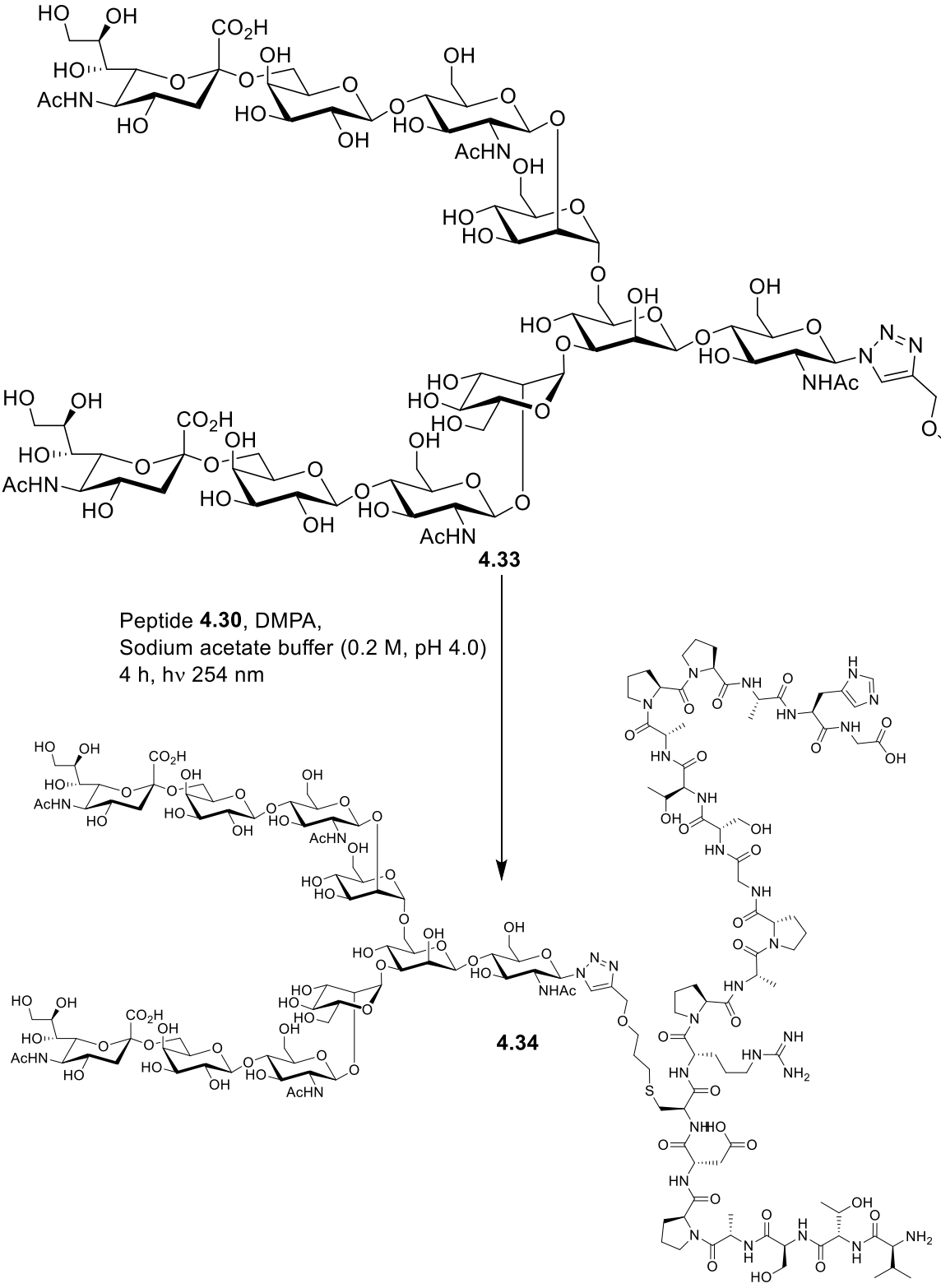


Scheme 4.18: Conjugation of complex biantennary glycan **4.32** to bifunctional linker **4.3**.

Having demonstrated that a monosaccharide could be successfully attached to a 20mer peptide, the next objective was to demonstrate that the sequential click reaction methodology could be applied to larger oligosaccharides. As previously (**Section 3.9**), the complex biantennary glycan **4.32** (**Scheme 4.18**) was selected as a model oligosaccharide substrate.

Sialoglycan **4.32** was prepared from hens' egg yolks[25, 26] by Dr Vivek Poonthiyil. Preparation of the glycosyl azide and subsequent attachment to linker **4.3** was carried out in a similar manner to all other conjugates prepared in this chapter (**Scheme 4.18**). As anticipated, a large excess of reagents was required to ensure complete conversion to glycosyl triazole **4.33**, which was isolated by HPLC in 94% yield.

However, attachment of sialoglycan alkene **4.33** to the modified MUC1 20mer peptide **4.30** (**Scheme 4.19**) proved challenging. The initial conditions selected (20 equivalents of sialoglycan-linker conjugate **4.33** and 0.1 equivalents of initiator DMPA with a reaction time of 4 h) resulted in no observable conversion to the desired product **4.34** (**Table 4.8**), as determined by HRMS. Increasing the concentration of reagents resulted in the formation of peptide disulfide, indicating that while radical initiation and chain transfer to thiol had occurred, recombination of the thiyl radicals produced was preferred over reaction with the linker alkene. The addition of TCEP to reduce any disulfide formed *in situ*[27, 28] reduced the peptide disulfide, but no product was observed.



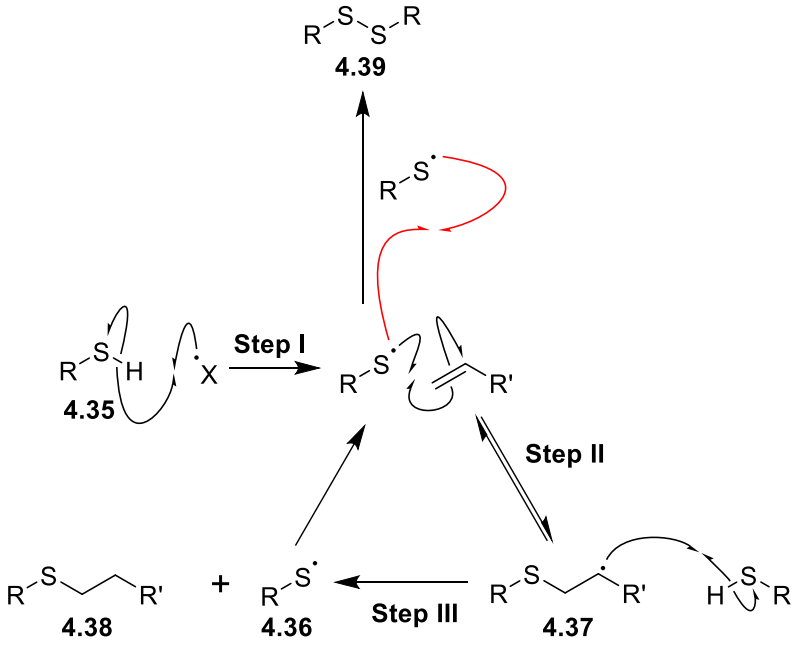
Scheme 4.19: Attempted attachment of sialoglycan alkene **4.33** to modified MUC1 20mer peptide **4.30**.

Concentration of 4.33 (mM)	Concentration of 4.31 (mM)	Concentration of DMPA (mM)	Concentration of TCEP (mM)	Result
12	0.6	0.06	0	No reaction
120	6	0.6	0	Peptide disulfide
120	6	0.6	60	No reaction
120	0.6	0.6	60	Trace **4.34**
120	0.6[a]	0.6	60	Trace **4.34**

Table 4.8: Thiol-ene reaction of sialoglycan-linker conjugate **4.33** with modified MUC1 peptide **4.30**. [a]Added over 4 h *via* syringe pump.

Increasing the number of equivalents of reagents used and lowering the concentration of peptide **4.31** used resulted in the formation of a trace amount of product, as observed by HRMS, presumably because the lower concentration of peptide **4.31** leads to a correspondingly lower thiyl radical concentration, and therefore the rate of radical recombination is reduced. It was reasoned that perhaps adding the peptide over an extended period of time *via* a syringe pump would lead to a further increase in product yield, but disappointingly this was not the case.

It was reasoned that perhaps the low conversion to product when both the sugar and peptide were large structures (as compared to previous examples in this chapter) was due to the slow hydrogen atom abstraction from another molecule of thiol by the thiol-ene conjugate (**Scheme 4.20**, **Step III**).



Scheme 4.20: Mechanism of the thiol-ene reaction.

The radical in the thiol-ene conjugate is between the bulky oligosaccharide and the large peptide (**4.37**), and must be approached by a sterically hindered peptide thiol (**4.35)** in order to abstract a hydrogen atom (**Step III**). **Step II** is reversible and therefore unless thiol-ene conjugate **4.37** abstracts a hydrogen atom it may decompose into its components. The rate at which thiyl radicals combine to form disulfide **4.39** is non-negligible, as demonstrated by the formation of peptide disulfide, and so presumably the reduced rate of **Step III** leads to the preferential formation of peptide disulfide.

With this hypothesis, methods for increasing the rate of hydrogen abstraction by the thiol-ene conjugate radical were explored. Of these, the most promising was the use of a radical chain transfer agent. The most well-known example of a radical chain transfer agent is tributyltin hydride.[29] However tin reagents are highly toxic and tin contaminants are notoriously difficult to remove.[30] One alternative to tributyltin hydride is tris(trimethylsilyl)silane (TTMSS).[31-33] While neither tributyltin hydride nor TTMSS have been studied in the context of thiol-ene click

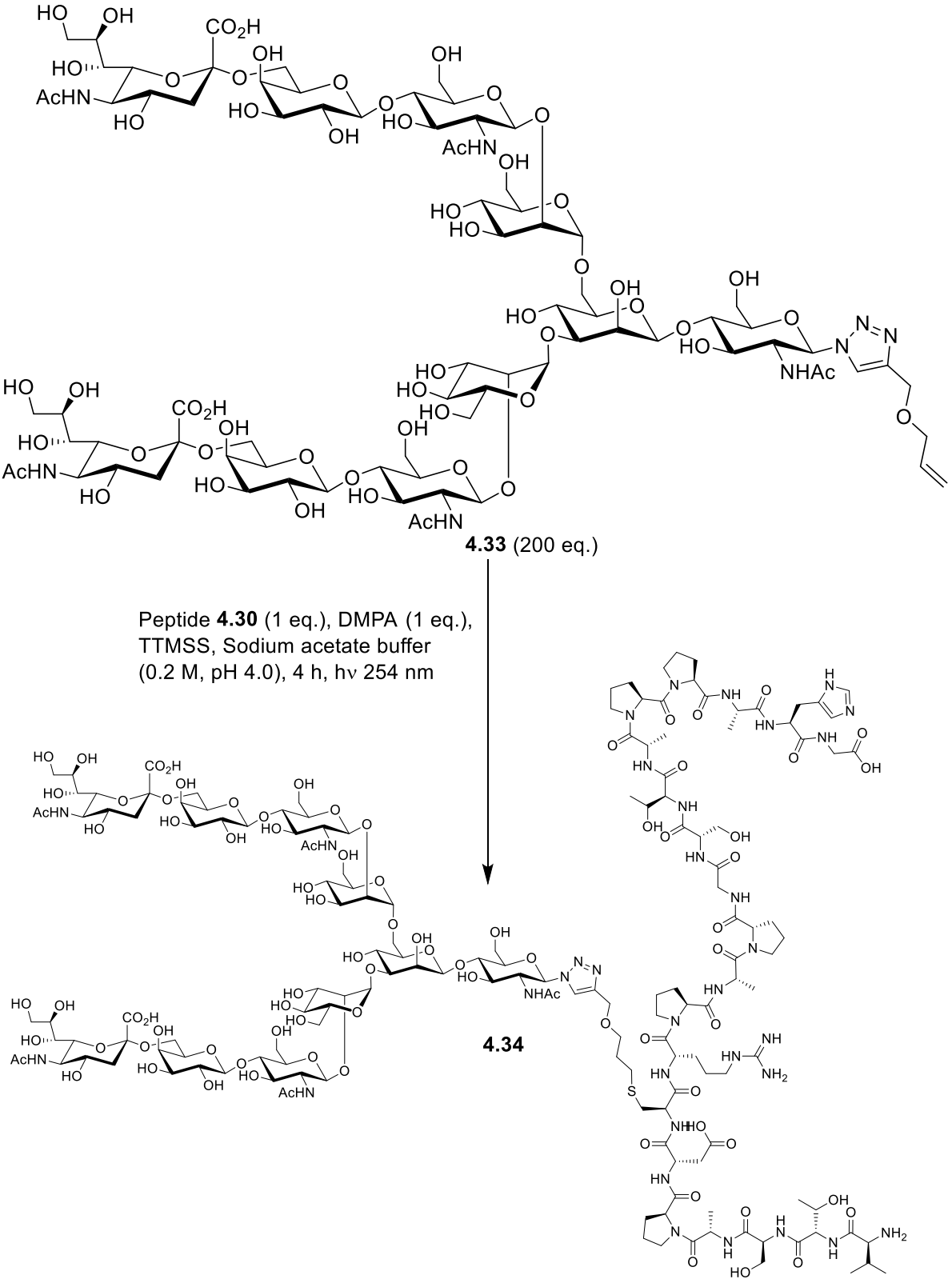
reactions, TTMSS has been demonstrated to act as a chain transfer agent in the addition of radicals to alkenes[32] and it has been demonstrated TTMSS can, under the correct conditions, abstract a hydrogen atom from a thiol.[34]

TTMSS equivalents relative to peptide 4.30	**Result**
0	Trace **4.34**
0.1	Trace **4.34**
1	Trace **4.34**

Table 4.9: Addition of TTMSS to the thiol-ene reaction between peptide **4.30** and sialoglycan alkene **4.33**.

Addition of TTMSS had no discernible effect on the thiol-ene click reaction between the modified MUC1 20mer peptide **4.30** and sialoglycan alkene **4.33** (**Scheme 4.21**, **Table 4.9**), implying either that TTMSS was not capable of donating a hydrogen atom to the peptide-oligosaccharide conjugate, that the TTMSS radical was not capable of abstracting a hydrogen atom from the cysteine residue of peptide **4.30** or that TTMSS directly reduced the thiol of peptide **4.30** too rapidly.

Due to the lack of progress in producing even a small quantity of sialoglycan glycoconjugate by this method, direct attachment of sialoglycan alkene **4.33** to peptide **4.30** was abandoned due to time constraints.

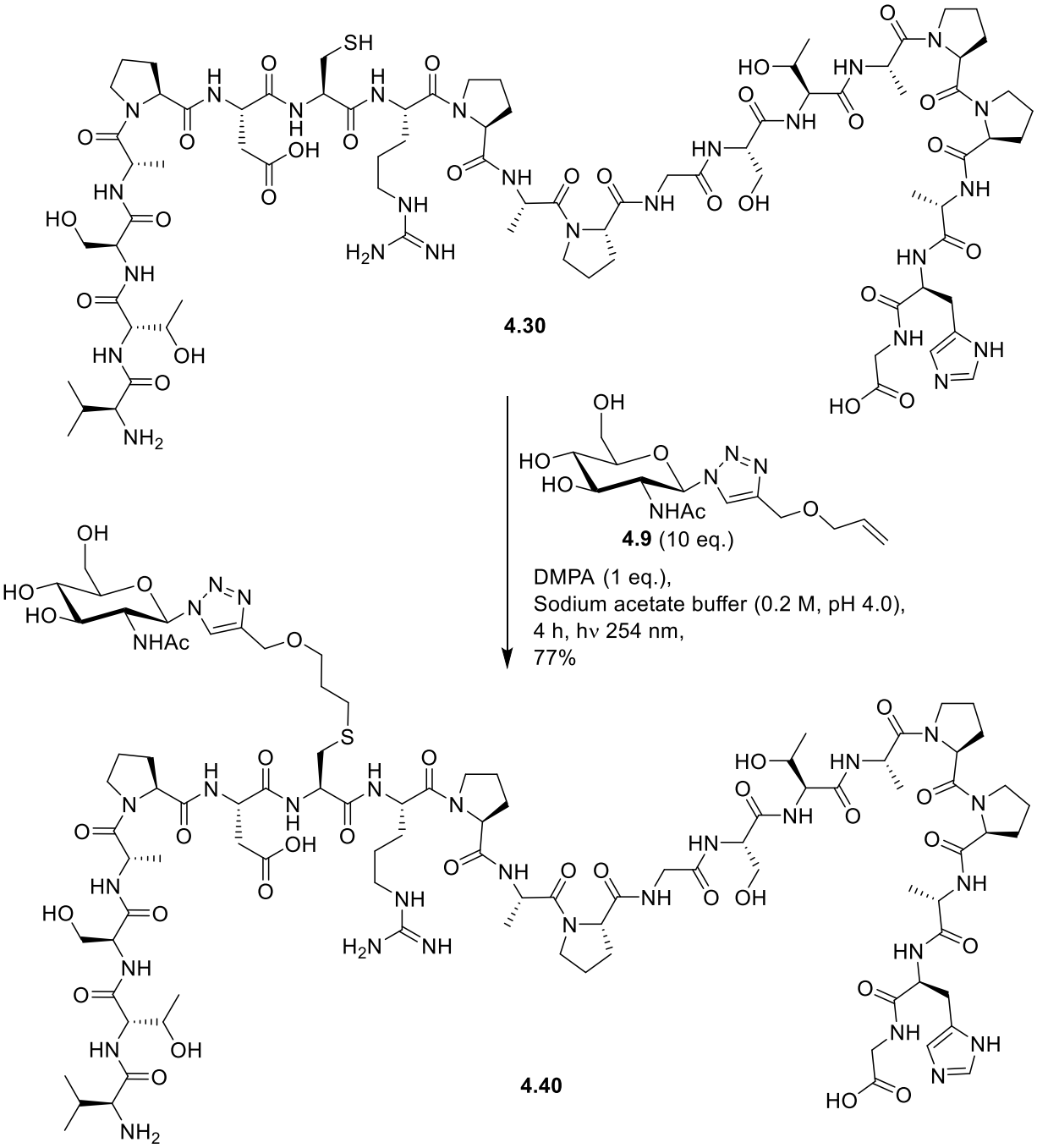


Scheme 4.21: Addition of TTMSS to the thiol-ene reaction between peptide **4.30** and sialoglycan-linker conjugate **4.33**.

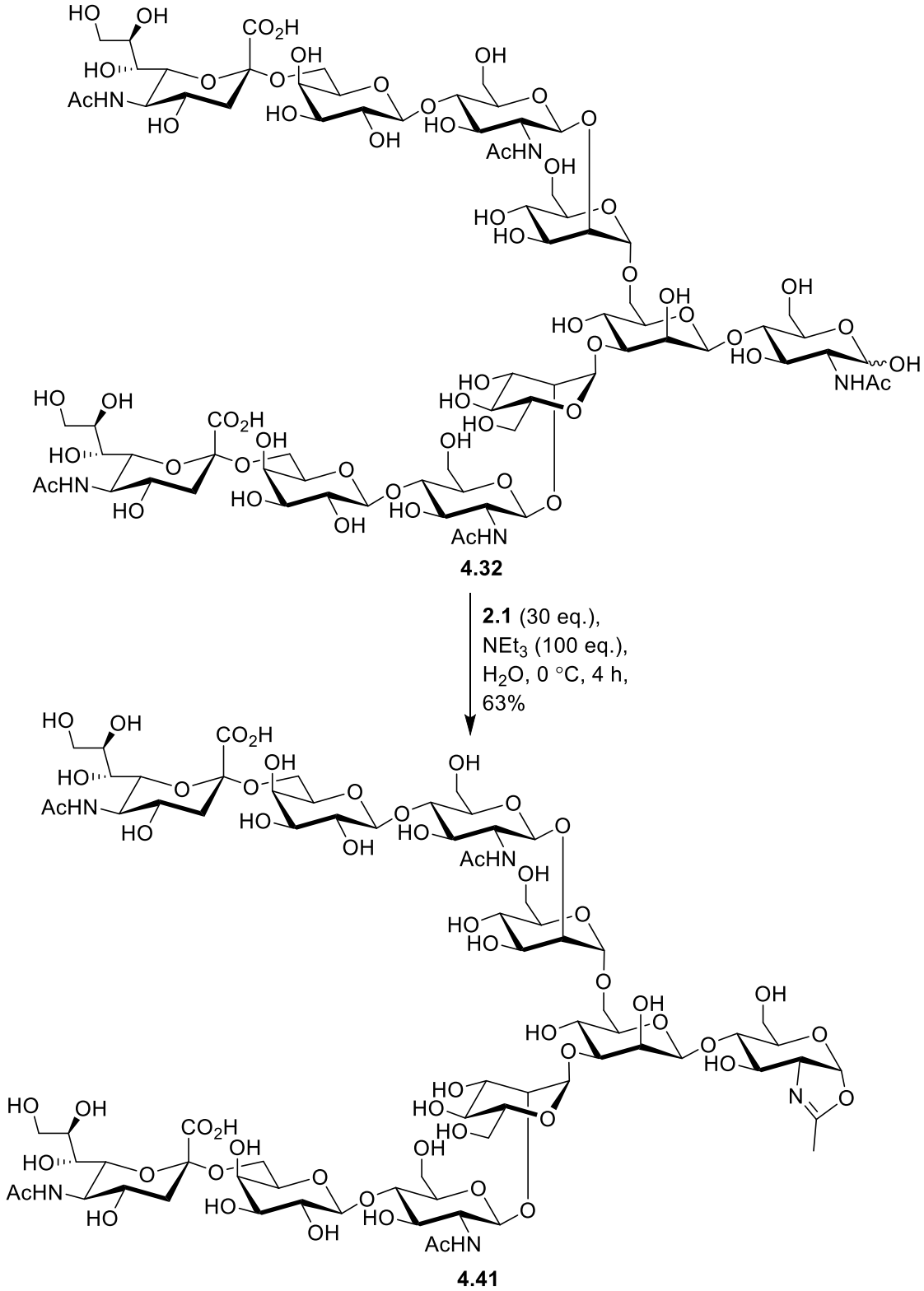
4.6.1 Sialoglycan attachment using a glycosynthase

While the direct attachment of a large oligosaccharide to modified MUC1 peptide **4.30** had not been possible, a sequence of reactions could be used to accomplish the same goal. Glycosynthase enzymes are modified glycosidase enzymes which are capable of catalysing the formation of a glycosidic bond.[35] Notably, Endo M N175Q, a mutant of Endo M, can catalyse the attachment of complex biantennary sialoglycan oxazoline to a wide variety of substrates, whilst it has no hydrolytic activity towards the product.[36, 37] If a peptide with a single GlcNAc attached could be prepared using two sequential click reactions then a sialoglycan oxazoline could be attached to the peptide using Endo M N175Q as a catalyst. In general, this method would allow the attachment of a GlcNAc "handle" to a peptide, which could then be further functionalised using mutant ENGase enzymes.

In order to test this proposal, GlcNAc alkene **4.9** was attached to modified MUC1 peptide **4.30** (**Scheme 4.22**). Using the same conditions that had been used for the attachment of GalNAc, the GlcNAc-peptide conjugate **4.40** was obtained in 77% yield, as determined by HPLC.



Scheme 4.22: Attachment of GlcNAc alkene **4.9** to the modified MUC1 20mer peptide **4.30**.

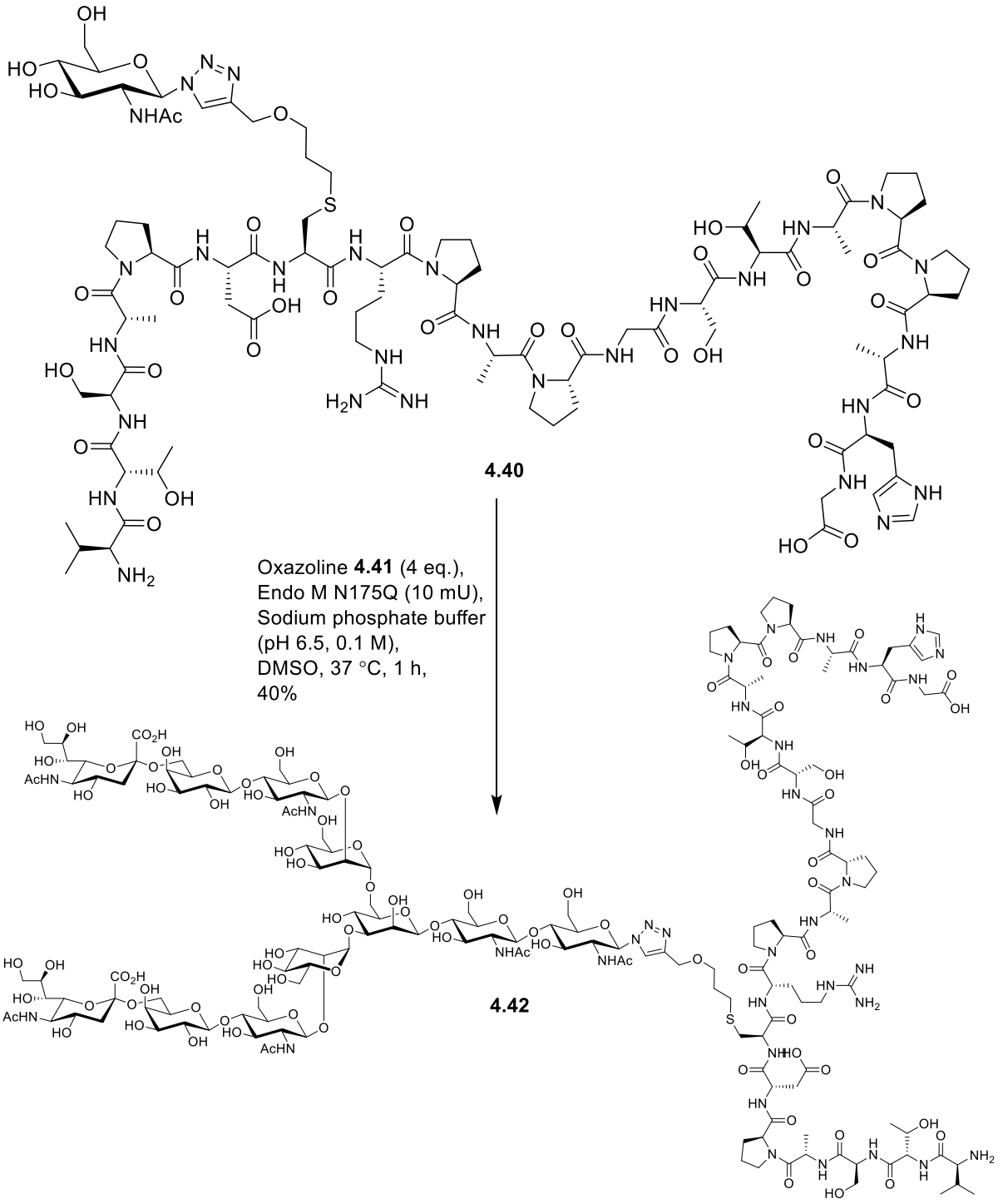


Scheme 4.23: Preparation of sialoglycan oxazoline **4.41**.

Sialoglycan oxazoline **4.41** was prepared using DMC **2.1** in a yield of 63% by following the literature procedure (**Scheme 4.23**).[38] A large excess of reagents was required to ensure complete reaction of sialoglycan **4.32**.

With both substrates in hand, the glycosynthase-catalysed glycosylation reaction of oxazoline **4.41** and GlcNAc conjugate **4.40** was examined. Pleasingly, the conditions reported by Tomabechi *et al*. (**Scheme 4.24**)[37] for the use of Endo M N175Q to catalyse similar reactions proved successful with only minor alterations (10 mU of Endo M N175Q was used instead of 2 mU), and afforded the desired glycopeptide **4.42** in a yield of 40% as determined by HPLC.

Thus, while it had not been possible to attach the large oligosaccharide directly to the MUC1 20mer through the use of two sequential click reactions using the conditions so far investigated, it was demonstrated that following the introduction of GlcNAc as a handle by double click chemistry the attachment of a larger *N*-glycan structure could be achieved by the use of an ENGase-catalysed reaction.



Scheme 4.24: Preparation of sialoglycan-modified MUC1 peptide glycopeptide **4.37**.

4.7 Conclusions

The use of a bifunctional linker to prepare glycoconjugates was demonstrated to be an effective means of attaching monosaccharide units and small oligosaccharide chains to a variety of substrates. Sugar-linker conjugates were prepared *via* CuAAC for a wide range of small sugars, and a number of different linkers. These sugar linker conjugates could then be attached to thiols *via* thiol-ene click chemistry. A small number of thiols were investigated, and it was determined that the effectiveness of the reaction largely hinged on the reactivity of the thiol. The reaction sequence could be applied to large, biologically relevant thiols, as demonstrated by the attachment of a GalNAc alkene to a modified MUC1 20mer peptide. However, while larger oligosaccharides could be attached to a linker, the subsequent thiol-ene click reaction proved challenging. In order to avoid the difficulties associated with this reaction, it was demonstrated that glycoconjugates containing larger oligosaccharides could be prepared by first attaching a GlcNAc residue to the peptide using two sequential click reactions, and then attaching by an oligosaccharide to the GlcNAc residue using an ENGase enzyme. While not being able to directly attach a large oligosaccharide is a current limitation of the methodology, the attachment of a GlcNAc handle followed by the use of a glycosynthase enzyme bypasses the problematic step.

Only a small subset of linkers were tested in this study. Therefore, in future it would be worthwhile to expand the set of linkers tested, with a view to increasing the effectiveness of the reaction sequence, particularly with larger substrates. Also, the linker structure differs significantly from natural glycoconjugate linkages, and therefore it would be beneficial to carry out studies of the biological activity of compounds produced in this manner, again investigating the effect of varying the linker.

4.8 References

1. van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G., *Nature* **2007,** *446* (7139), 1105-1109.

2. Fernandez-Gonzalez, M.; Boutureira, O.; Bernardes, G. J. L.; Chalker, J. M.; Young, M. A.; Errey, J. C.; Davis, B. G., *Chem. Sci.* **2010,** *1* (6), 709-715.

3. Tanaka, T.; Inoue, G.; Shoda, S.-I.; Kimura, Y., *J. Polym. Sci., Part A: Polym. Chem.* **2014,** *52* (24), 3513-3520.

4. Tanaka, T.; Ishitani, H.; Miura, Y.; Oishi, K.; Takahashi, T.; Suzuki, T.; Shoda, S. I.; Kimura, Y., *ACS Macro Letters* **2014,** *3* (10), 1074-1078.

5. Tanaka, T.; Matsumoto, T.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Lett.* **2009,** *38* (5), 458-459.

6. Novoa, A.; Barluenga, S.; Serba, C.; Winssinger, N., *Chem. Commun.* **2013,** *49* (69), 7608-7610.

7. Dondoni, A.; Marra, A., *Chem. Soc. Rev.* **2012,** *41* (2), 573-586.

8. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J., *Angew. Chem., Int. Ed.* **2014,** *53* (44), 1907-11911.

9. McSweeney, L.; Dénès, F.; Scanlan, E. M., *Eur. J. Org. Chem.* **2016,** *2016* (12), 2080-2095.

10. Dadová, J.; Orság, P.; Pohl, R.; Brázdová, M.; Fojta, M.; Hocek, M., *Angew. Chem., Int. Ed.* **2013,** *52* (40), 10515-10518.

11. Zhou, K.; Tong, L.; Deng, J.; Yang, W., *J. Mater. Chem.* **2010,** *20* (4), 781-789.

12. Maleczka Jr, R. E.; Ghosh, B.; Gallagher, W. P.; Baker, A. J.; Muchnij, J. A.; Szymanski, A. L., *Tetrahedron* **2013,** *69* (19), 4000-4008.

13. Trost, B. M.; Xie, J., *J. Am. Chem. Soc.* **2006,** *128* (18), 6044-6045.

14. Wilkinson, B. L.; Bornaghi, L. F.; Poulsen, S.-A.; Houston, T. A., *Tetrahedron* **2006,** *62* (34), 8115-8125.

15. Dedola, S.; Hughes, D. L.; Nepogodiev, S. A.; Rejzek, M.; Field, R. A., *Carbohydr. Res.* **2010,** *345* (9), 1123-1134.

16. Worrell, B. T.; Malik, J. A.; Fokin, V. V., *Science* **2013,** *340* (6131), 457-460.

17. Tanaka, T.; Nagai, H.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Commun.* **2009,** (23), 3378-3379.

18. Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G., *Angew. Chem., Int. Ed.* **2009,** *48* (42), 7798-7802.

19. Lal, M.; Rao, R.; Fang, X.; Schuchmann, H.-P.; von Sonntag, C., *J. Am. Chem. Soc.* **1997,** *119* (24), 5735-5739.

20. Hanisch, F.-G.; Müller, S., *Glycobiology* **2000,** *10* (5), 439-449.

21. Taylor-Papadimitriou, J.; Burchell, J.; Miles, D. W.; Dalziel, M., *Biochim. Biophys. Acta Mol. Bas. Dis.* **1999,** *1455* (2–3), 301-313.

22. von Mensdorff-Pouilly, S.; Petrakou, E.; Kenemans, P.; van Uffelen, K.; Verstraeten, A. A.; Snijdewint, F. G. M.; van Kamp, G. J.; Schol, D. J.; Reis, C. A.; Price, M. R.; Livingston, P. O.; Hilgers, J., *Int. J. Cancer* **2000,** *86* (5), 702-712.

23. Gendler, S.; Taylor-Papadimitriou, J.; Duhig, T.; Rothbard, J.; Burchell, J., *J. Biol. Chem.* **1988,** *263* (26), 12820-12823.

24. Zhang, S.; Zhang, H. S.; Cordon-Cardo, C.; Reuter, V. E.; Singhal, A. K.; Lloyd, K. O.; Livingston, P. O., *Int. J. Cancer* **1997,** *73* (1), 50-56.

25. Seko, A.; Koketsu, M.; Nishizono, M.; Enoki, Y.; Ibrahim, H. R.; Juneja, L. R.; Kim, M.; Yamamoto, T., *Biochim. Biophys. Acta Gen. Subjects* **1997,** *1335* (1), 23-32.

26. Sun, B.; Bao, W.; Tian, X.; Li, M.; Liu, H.; Dong, J.; Huang, W., *Carbohydr. Res.* **2014,** *396*, 62-69.

27. Healy, J.; Ekkerman, S.; Pliotas, C.; Richard, M.; Bartlett, W.; Grayer, S. C.; Morris, G. M.; Miller, S.; Booth, I. R.; Conway, S. J., *Biochemistry* **2014,** *53* (12), 1982-1992.

28. Hoppmann, C.; Kühne, R.; Beyermann, M., *Beilstein J. Org. Chem.* **2012,** *8*, 884.

29. Curran, D. P., *Synthesis* **1988,** *1988* (06), 417-439.

30. Le Grognec, E.; Chrétien, J.-M.; Zammattio, F.; Quintard, J.-P., *Chem. Rev.* **2015,** *115* (18), 10207-10260.

31. Chatgilialoglu, C.; Guerra, M.; Guerrini, A.; Seconi, G.; Clark, K.; Griller, D.; Kanabus-Kaminska, J.; Martinho-Simoes, J., *J. Org. Chem.* **1992,** *57* (8), 2427-2433.

32. Giese, B.; Kopping, B., *Tetrahedron Lett.* **1989,** *30* (6), 681-684.

33. Chatgilialoglu, C.; Lalevée, J., *Molecules* **2012,** *17* (1), 527-555.

34. Chatgilialoglu, C., *Helv. Chim. Acta* **2006,** *89* (10), 2387-2398.

35. Fairbanks, A. J. The ENGases: versatile biocatalysts for the production of homogeneous N-linked glycopeptides and glycoproteins. *Chem. Soc. Rev.* [Online early access]. DOI: 10.1039/C6CS00897F. Published online: Jul 6, 2017. http://dx.doi.org/10.1039/C6CS00897F (accessed Jul 11, 2017).

36. Umekawa, M.; Li, C.; Higashiyama, T.; Huang, W.; Ashida, H.; Yamamoto, K.; Wang, L.-X., *J. Biol. Chem.* **2010,** *285* (1), 511-521.

37. Tomabechi, Y.; Squire, M. A.; Fairbanks, A. J., *Org. Biomol. Chem.* **2014,** *12* (6), 942-955.

38. Noguchi, M.; Tanaka, T.; Gyakushi, H.; Kobayashi, A.; Shoda, S. I., *J. Org. Chem.* **2009,** *74* (5), 2210-2212.

Chapter 5: Flavinium salt-catalysed synthesis of glycosyl sulfoxides and investigation of their activity as glycosidase inhibitors

5.1 Flavinium salts

In recent years there has been an intense focus on "green chemistry",[1] developing reactions which are less harmful to the environment. The application of flavinium salts (**Section 1.14.1**) appears to represent a real possibility to improve the synthesis of glycosyl sulfoxides, as it has been demonstrated to be clean, producing only water as a byproduct, and efficient, oxidising a wide range of sulfides to sulfoxides.[2-4] The use of hydrogen peroxide as an oxidant is desirable because it is both atom-economic,[5] and produces only environmentally benign water as a byproduct, while the use of water as a solvent is preferred over organic solvents because it is environmentally benign.

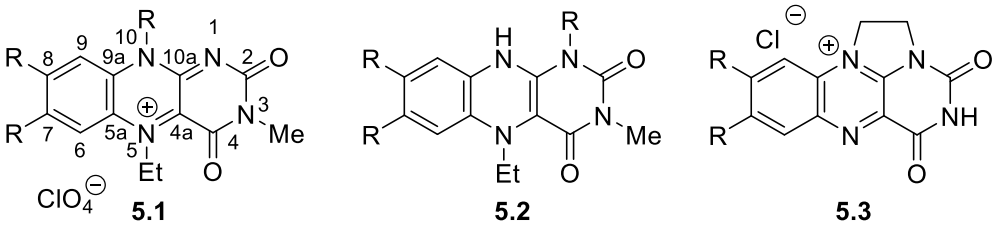
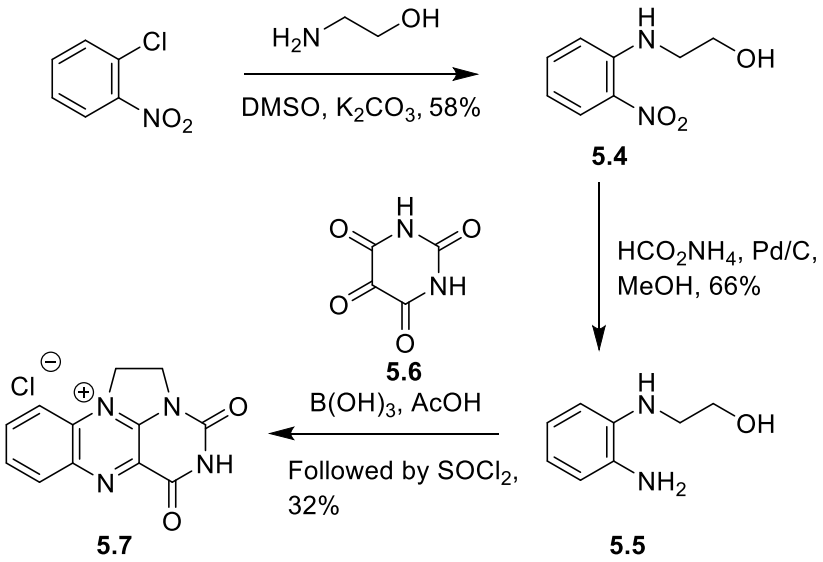


Figure 5.1: Flavin species that have previously been examined as catalysts for the oxidation of sulfides.

The activities of the various types of flavin species (**Figure 5.1**), N^5,N^{10} substituted flavinium salts (**5.1**),[2] N^1,N^5 substituted dihydroalloxazines (**5.2**),[3] and N^1,N^{10} bridged flavinium salts (**5.3**),[4] are broadly comparable, although direct comparisons are scarce. Therefore, N^1,N^{10} substituted flavinium salts were selected for testing as catalysts for the oxidation of thioglycosides to glycosyl sulfoxides, as their preparation is simpler and shorter, a beneficial

feature if such catalysts are to be used on a large scale. Furthermore, they are easily tuneable, notably by modification of either the aromatic ring or the bridging chain,[4, 6] and they have been shown to be more stable than some of the alternatives, notably N^1,N^5 substituted flavinium salts.[3]

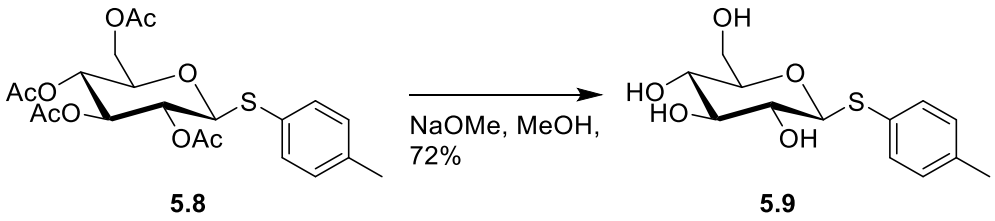
5.2 Preparation of catalyst 5.7



Scheme 5.1: Preparation of flavinium salt catalyst **5.7**.

The N^1,N^{10} bridged flavinium salt **5.7** was prepared according to the literature procedure (**Scheme 5.1**).[7] 1-Chloro-2-nitrobenzene was converted to compound **5.4** *via* nucleophilic aromatic substitution by ethanolamine, using potassium carbonate as a base. The nitro- group of **5.4** was converted to an amine by catalytic reduction, with palladium on carbon as a catalyst and ammonium formate as the hydrogen source. Condensation of diamine **5.5** with alloxan **5.6** catalysed by boric acid and acetic acid, followed by conversion of the free hydroxyl group to a chloride using thionyl chloride and then *in-situ* nucleophilic substitution of the chloride by N^1 afforded catalyst **5.7**.

5.3 Preparation of the test substrate



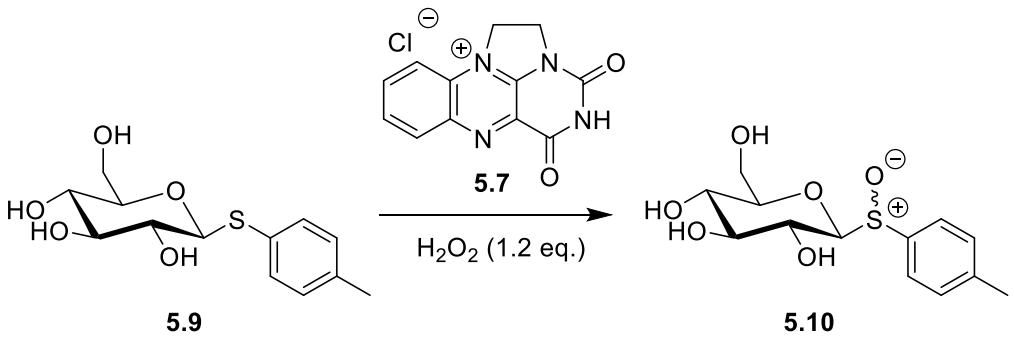
5.8 **5.9**

Scheme 5.2: Preparation of tolyl thioglycoside **5.9**.

Tolyl thioglycoside **5.9** was chosen as a test substrate for the oxidation reaction, as it is unprotected and therefore expected to be more easily oxidised in an aqueous environment than protected sulfoxides. Thioglycoside **5.9** was prepared by Zemplén deacetylation[8] of the known acetyl-protected tolyl thioglycoside **5.8**[9] (**Scheme 5.2**), which had previously been synthesised by another member of the Fairbanks group.

5.4 Investigation of the oxidation reaction

With catalyst **5.7** and model substrate **5.9** in hand, investigation of the oxidation reaction commenced (**Scheme 5.3**). A loading of 1.8% of catalyst **5.7** was used, as this loading has previously been shown to catalyse oxidation reactions effectively.[4] Only 1.2 equivalents of hydrogen peroxide were used, as peroxide is potentially explosive and therefore minimising the amount used is desirable. Screening of reaction solvents revealed that methanol was most effective, the use of which resulted in 78% conversion to product as determined by ^{1}H NMR spectroscopy (**Table 5.1**). The use of acetonitrile resulted in the production of significant quantities of impurities. Water would be an ideal solvent for this reaction, as it is environmentally benign and readily available, and would therefore further complement the conditions required. Unfortunately tolyl thioglycoside **5.9** was not soluble in water. Methanol was therefore arrived upon as the optimal solvent for the oxidation of tolyl thioglycoside **5.9**.



Scheme 5.3: Oxidation of tolyl thioglycoside **5.9** to glycosyl sulfoxide **5.10** catalysed by flavinium salt **5.7**.

Solvent	Catalyst Loading	Reaction time (h)	Conversion
H_2O	1.8%	6	0%
MeCN	1.8%	6	Impurities
MeOH	1.8%	6	78%
MeOH	5%	6	87%
MeOH	10%	6	84%
MeOH	20%	6	92%
MeOH	5%	4	69%
MeOH	5%	8	86%

Table 5.1: Optimisation of oxidation of tolyl glycoside **5.9** to glycosyl sulfoxide **5.10**.

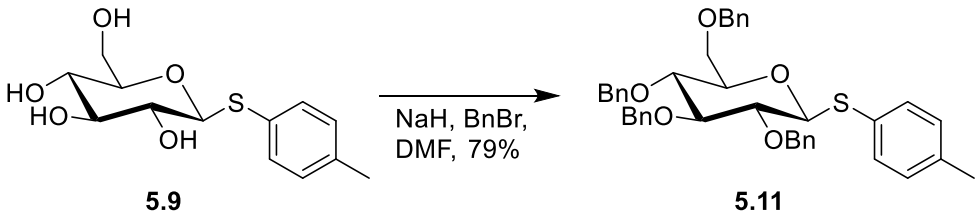
Next, other reaction parameters were systematically varied. Increasing the catalyst loading resulted in a small increase in conversion. At a loading of 5% catalyst 87% conversion to product was observed; an increase of 9% over the oxidation reaction with a catalyst loading of 1.8%. Increasing the catalyst loading to 10% resulted in 84% conversion to product, and further increasing the catalyst loading to 20% gave a conversion of 92%. However, the increase in conversion upon increasing the amount of catalyst above 5% loading was slight, and appears to be within the experimental margin of error. Therefore, 5% was selected as a suitable catalyst

loading for further optimisation and exemplification. Six hours was found was found to be the optimum reaction time; reactions run for less time had significantly lower conversions, whilst the use of more extended reaction times resulted in no significant change in conversion.

Purification of sulfoxide **5.10** proved challenging, as it is insoluble in a number of common chromatography solvents, non-crystalline, and unstable to both acid and heat. Purification was eventually accomplished with a yield of 74% by pre-absorption of **5.10** onto Florisil®, followed by rapid flash column chromatography with CH_2Cl_2:MeOH (7:1) as the solvent.

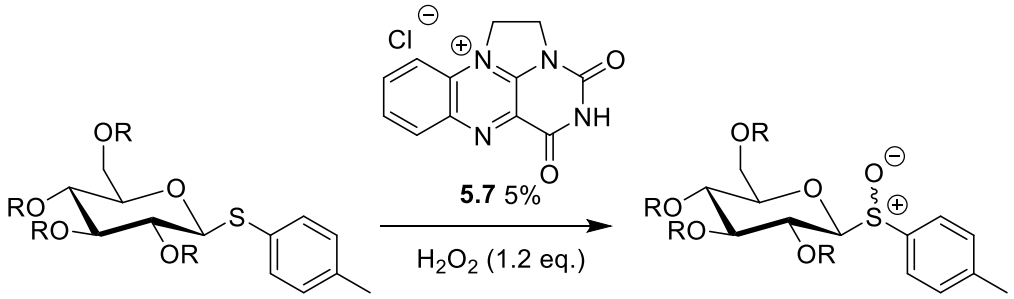
5.5 Oxidation of protected thioglycosides

Having demonstrated that flavin **5.7** is an effective catalyst for the hydrogen peroxide-mediated oxidation of unprotected tolyl thioglycosides, attention turned to protected thioglycosides. Benzylated tolyl thioglycoside **5.11**[9] was prepared using standard conditions, by reaction of unprotected tolyl thioglycoside **5.9** with benzyl bromide, in the presence of sodium hydride as a base (**Scheme 5.4**). Acetylated tolyl thioglycoside **5.8**[9] was already available, and together these protecting groups represent those most commonly used in carbohydrate chemistry.



5.9 **5.11**

Scheme 5.4: Preparation of benzylated tolyl thioglycoside **5.11**.



5.8 R = Ac
5.11 R = Bn

5.12 R = Ac
5.13 R = Bn

Scheme 5.5: Oxidation of protected tolyl thioglycosides catalysed by flavinium salt **5.7**.

Compound	Solvent	Temperature	Reaction Time (h)	Conversion
5.8	H_2O	RT	6	-
5.8	MeOH	RT	6	-
5.8	MeOH	Reflux	6	22%
5.8	DCM	RT	6	-
5.8	MeCN	RT	6	30%
5.8	MeCN	Reflux	6	Decomp.
5.8	MeCN	RT	72	45%
5.11	MeCN	RT	6	<10%

Table 5.2: Testing of the hydrogen peroxide-mediated oxidation of protected tolyl thioglycosides catalysed by flavinium salt **5.7**.

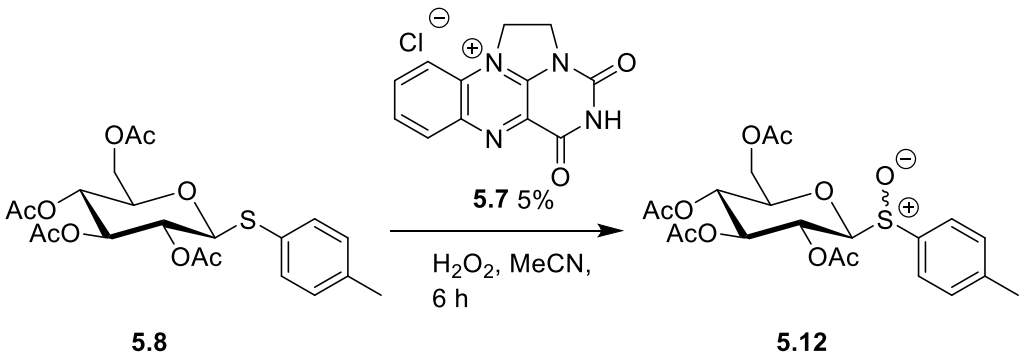
First, the conditions established for unprotected tolyl thioglycoside **5.9** were applied to acetylated tolyl thioglycoside **5.8** (**Scheme 5.5**). Unfortunately, it quickly became apparent that the reaction was significantly less effective with protected thioglycosides (**Table 5.2**). Acetylated thioglycoside **5.8** is insoluble in water and only sparingly soluble in methanol, and therefore perhaps unsurprisingly no reaction was observed in either of these solvents. Heating the reaction to reflux in methanol increased the rate of reaction, but after six hours only 22%

conversion to **5.12** was observed. The use of dichloromethane as a solvent again resulted in no reaction. Both catalyst **5.7** and hydrogen peroxide are insoluble in dichloromethane, leading to a heterogeneous system. This is consistent with previous studies of N^1,N^{10} bridged flavinium catalysts, where it was reported that the use of dichloromethane as a reaction solvent resulted in irreproducible results due to heterogeneity of the reaction mixture.[4]

The use of acetonitrile as a reaction solvent resulted in a more promising 30% conversion of **5.8** into **5.12** after six hours at room temperature. As heating the reaction to reflux increased the rate of the corresponding reaction in methanol, refluxing the acetonitrile reaction was attempted. Unfortunately, analysis of the reaction mixture showed that a complex mixture of products had been formed. Sulfoxides are known to be unstable when heated, decomposing into highly reactive sulfenic acids.[10, 11] Indeed, in the course of this work even heating glycosyl sulfoxides *in vacuo* while concentrating the crude reaction mixture was observed to be sufficient to result in violent decomposition.

Extending the reaction time did result in an increase in conversion. After 72 hours only 45% conversion to product was observed, which, while an improvement, is far from ideal.

Application of identical conditions to the benzylated tolyl thioglycoside **5.11** gave a reduced conversion to the corresponding product **5.13**. Further optimisation attempts were therefore focused on acetylated thioglycoside **5.8**.



Scheme 5.6: Optimisation of the oxidation of acetylated tolyl thioglycoside **5.8**.

Loading of 5.7	**H_2O_2 equivalents**	**Conversion**
5%	1.2	30%
10%	1.2	39%
20%	1.2	39%
5%	2.4	63%
5%	7.2	94%

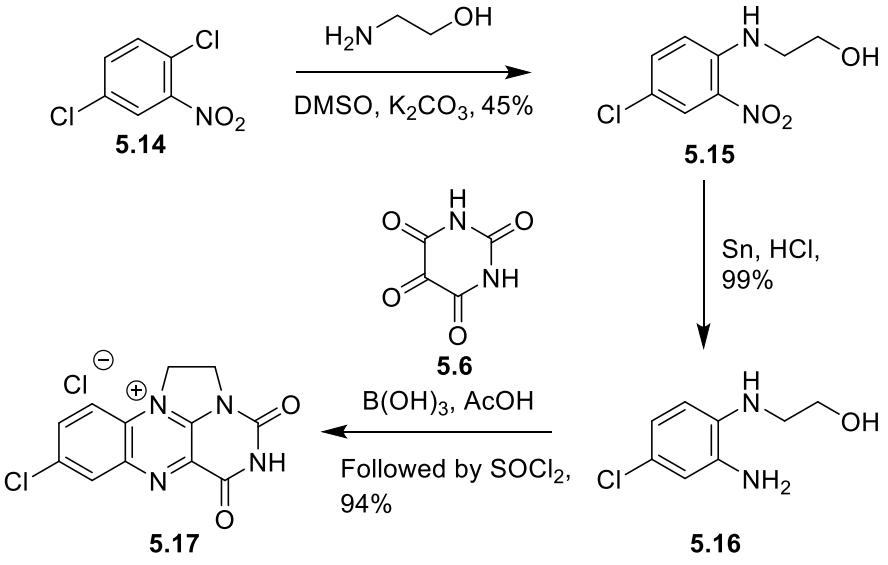
Table 5.3: Variation of reaction conversion with catalyst loading and the quantity of hydrogen peroxide for oxidation of acetylated tolyl thioglycoside **5.8**.

Due to the significantly lower conversion to product observed for protected tolyl thioglycosides, the catalyst loading and hydrogen peroxide concentration were reinvestigated (**Table 5.3**). Increasing the loading of catalyst **5.7** resulted in only a small increase in conversion, consistent with previous results. However, increasing the number of equivalents of hydrogen peroxide used did result in a more effective reaction, increasing the conversion of thioglycoside **5.8** to sulfoxide **5.12** to 94% when 7.2 equivalents of peroxide were used. However, given that hydrogen peroxide is a potentially explosive compound, and furthermore that the atom economy benefits of using peroxide as an oxidant are diminished when 7.2 equivalents are required, existing peroxide oxidation procedures, such as those catalysed by

silica,[12] were deemed more effective and benign. Therefore, this protocol was not further applied to protected sugars, as it was not judged to be an improvement on existing methodology.

5.5.1 Modification of the flavinium catalyst

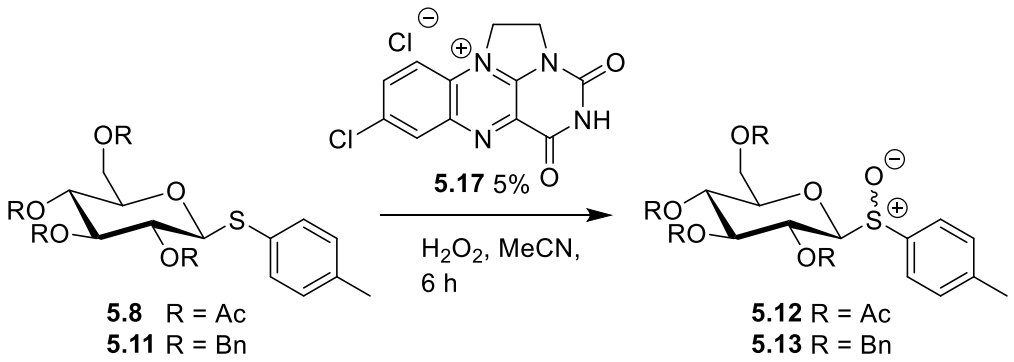
Unlike the reactions of N^1,N^5 substituted and N^5,N^{10} substituted flavinium salts[13-15] there has been no thorough mechanistic study of oxidation reactions catalysed by N^1,N^{10} bridged flavinium salts. However, electron-withdrawing subsituents attached to C7 and/or C8 have been shown to increase their catalytic activity,[4] implying that the rate-determining step is the elimination of water from the 10a adduct to regenerate the flavinium salt. Variation of these substituents could therefore provide a route to enhancing the oxidation of protected thioglycosides.



Scheme 5.7: Preparation of chlorine-substituted catalyst **5.17**.[7]

With the aim of testing this hypothesis, the chlorinated flavinium salt **5.17** was prepared according to the literature procedure.[7] An analogous reaction sequence to that used previously

was applied to 1,4-dichloro-2-nitrobenzene **5.14**. Regioselective nucleophilic aromatic substitution was carried out with ethanolamine using potassium carbonate as a base to obtain compound **5.15**. Reduction of the nitro group to afford diamine **5.16** was carried out using mossy tin[16] and hydrochloric acid in order to avoid dehalogenation. Diamine **5.16** was then condensed with alloxan **5.6** using boric acid as a catalyst. Ring-closure by conversion of the free hydroxyl group to a chloride followed by nucleophilic substitution afforded flavinium salt **5.17**.



Scheme 5.8: Oxidation of protected tolyl thioglycosides catalysed by flavinium salt **5.17**.

Compound	H_2O_2 equivalents	Conversion
5.8	1.2	38%
5.8	2.4	42%
5.8	3.6	47%
5.11	1.2	<10%

Table 5.4: Oxidation of protected tolyl thioglycosides catalysed by chlorinated flavinium salt **5.17**.

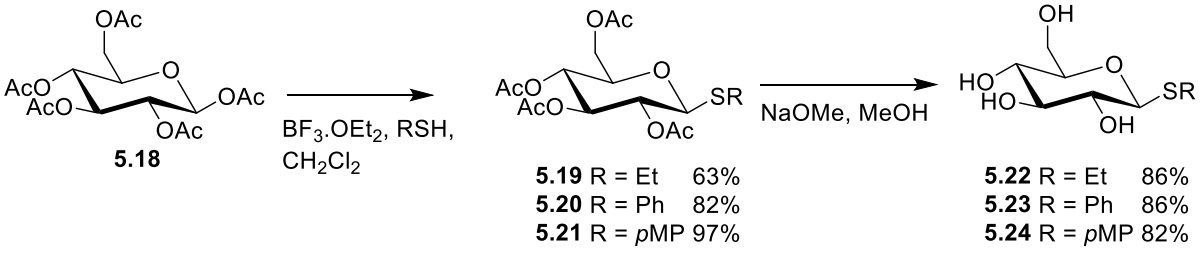
Testing of flavinium salt **5.17** did reveal a small increase in conversion of the acetylated thioglycoside **5.8** to the sulfoxide **5.12** (**Scheme 5.8**, **Table 5.4**). However, even the use of 3.6 equivalents of hydrogen peroxide only resulted in 47% conversion, significantly below existing

procedures. Similarly, no significant improvement for the conversion of benzylated tolyl thioglycoside **5.11** to sulfoxide **5.13** was observed.

With these results in mind, further investigation into the use of flavinium catalysts to catalyse the oxidation of protected thioglycosides were discontinued, as no conditions that were an improvement over pre-existing methods had been found. It was believed that solubility effects contributed to the poor performance of protected sugars in the flavinium salt catalysed reaction; in most of the solvents tested at least one of the protected sugar, hydrogen peroxide, or the flavinium salt catalyst was insoluble, or only partially soluble.

5.6 Substrate scope: variation of the aglycon

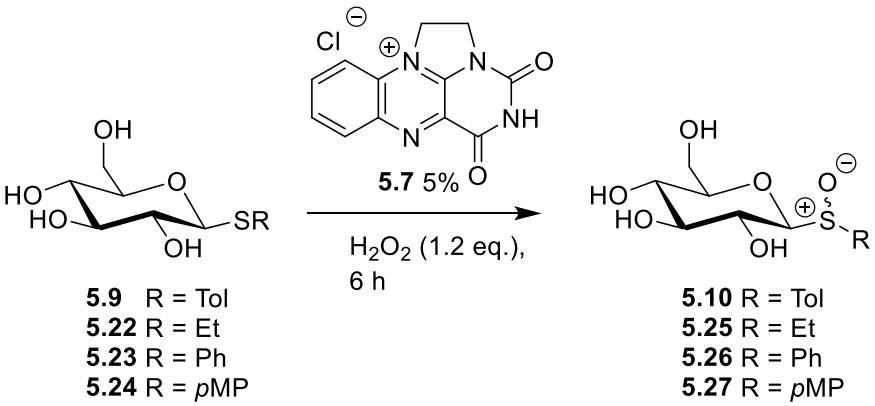
While protected sugars are required for the majority of applications of glycosyl sulfoxides, such as their use as glycosyl donors, there are some exceptions. Notably, glycosidase inhibitors possess free hydroxyl groups, as the natural substrates of these enzymes are themselves unprotected. Therefore, the substrate scope of the reaction of unprotected thioglycosides was investigated.



Scheme 5.9: Preparation of unprotected thioglucosides **5.22**-**5.24**.

Firstly, the aglycon was varied. A selection of common thioglucosides were prepared following the standard procedure (**Scheme 5.9**).[17] Peracetyl-β-glucose **5.18**[18] was converted to the

corresponding ethyl, phenyl, and *para*-methoxyphenyl thioglucosides **5.19**-**5.21** by activation with boron trifluoride in the presence of the relevant thiol. These reactions afforded protected ethyl-,[18] phenyl-,[19] and *para*-methoxyphenyl-[20] thioglucosides (**5.19**-**5.21**) in moderate to high yields. Deacetylation was then carried out under Zemplén conditions to afford the unprotected thioglucosides **5.22**-**5.24**[19-21] in high yields.



Scheme 5.10: Oxidation of various unprotected thioglycosides to glycosyl sulfoxides catalysed by flavinium salt **5.7**.

Compound	Solvent	Conversion	Isolated yield
5.9	MeOH	87%	74%
5.22	MeOH	93%	76%
5.22	H_2O	>99%	71%
5.23	MeOH	30%	-
5.24	MeOH	64%	-

Table 5.5: Oxidation of various unprotected thioglucosides to glycosyl sulfoxides catalysed by **5.3**.

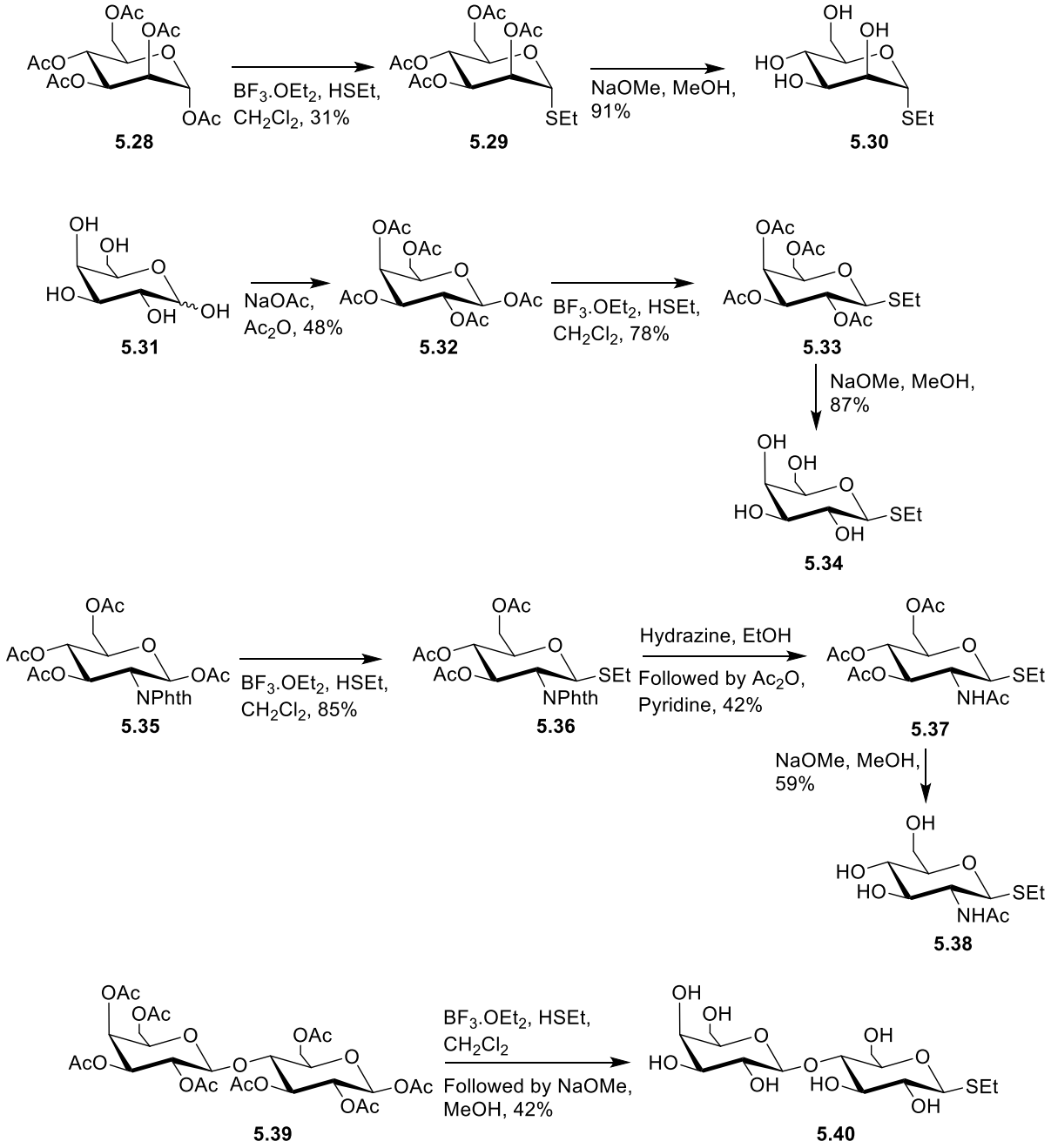
Testing revealed that ethyl thioglycoside **5.22** (93% conversion) was slightly more efficiently oxidised to sulfoxide than tolyl thioglycoside **5.9** (87% conversion). On the other hand, *para*-

methoxyphenyl thioglycoside **5.24** was a less effective substrate (64% conversion) and phenyl thioglycoside **5.23** was a poor substrate with only a very low 30% conversion to sulfoxide. Only tolyl glycosyl sulfoxide **5.10** and ethyl glycosyl sulfoxide **5.25** were isolable, due to the difficulty separating the unreacted thioglycoside starting material from the glycosyl sulfoxide. Again, purification proved challenging due to the high polarity and relative instability of the glycosyl sulfoxides when exposed to silica, but was accomplished by pre-absorbing the sulfoxides onto Florisil® before purifying by rapid flash column chromatography.

Interestingly, both the more electron-rich (*para*-methoxyphenyl- as compared to tolyl- groups) and the less electron-rich (phenyl- as compared to tolyl- groups) substrates were found to be less effective substrates than the tolyl thioglycoside. The precise reasons for this were not investigated, but one possible explanation is as follows: if a more electron-rich species reacts at a reduced rate then it is logical that the increased build-up of negative charge on the sulfur atom interferes with its ability to attack the flavinium peroxide species. On the other hand, for less electron-rich species attack on the flavinium peroxide species becomes more challenging due to the greatly reduced electron density at sulfur.

It was observed that ethyl thioglucoside **5.22** is soluble in water, unlike tolyl thioglucoside **5.9**. The efficiency of the reaction in water as the solvent was therefore investigated for ethyl thioglucoside **5.22**. Pleasingly, the use of water as a solvent led to the complete conversion of **5.22** to sulfoxide **5.25** (as observed by ^{1}H NMR spectroscopy). As water is the solvent of choice for this reaction, and also gives higher conversions in this case, ethyl thioglycosides were used for the remaining studies.

5.7 Substrate scope: variation of the sugar



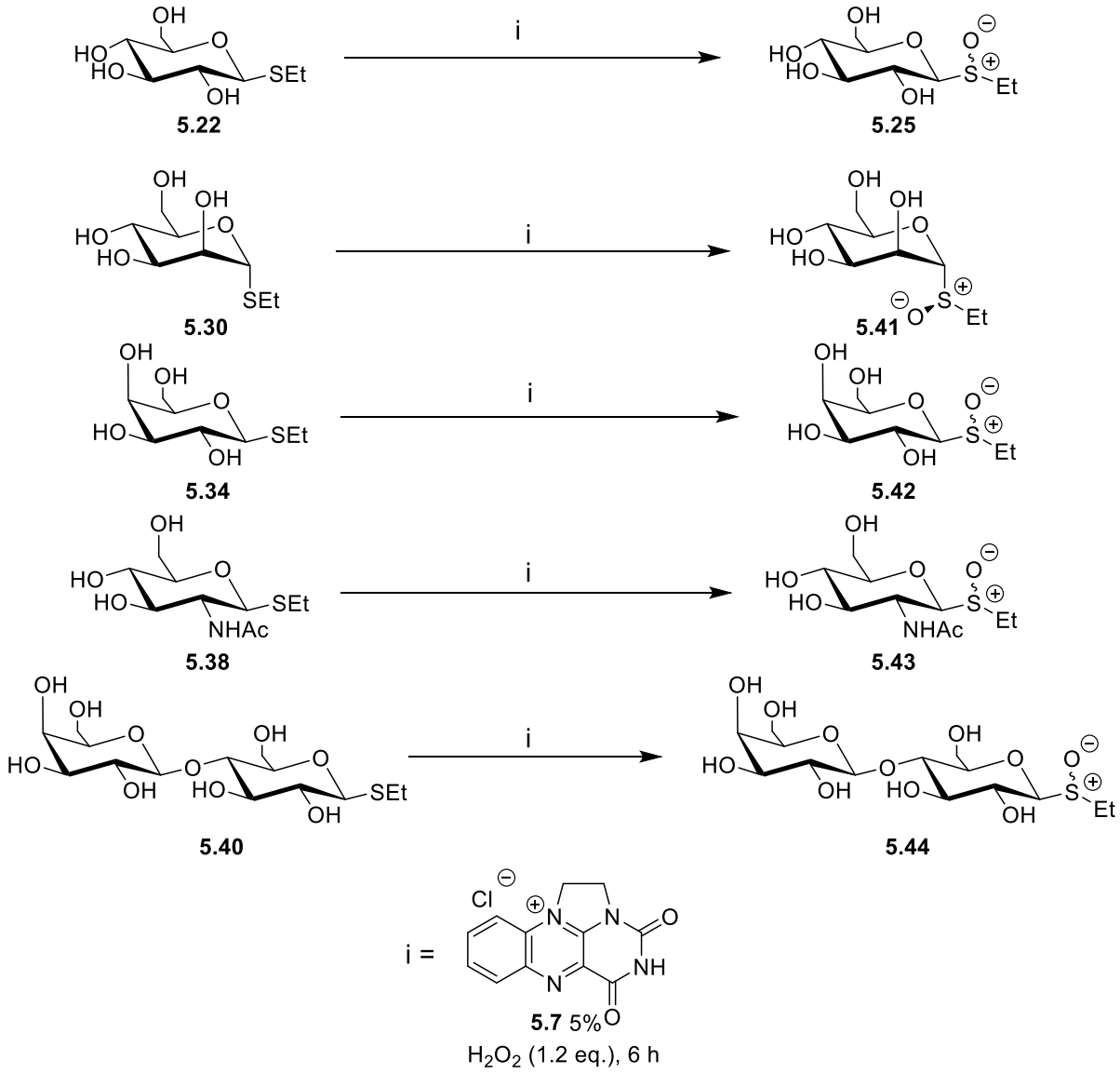
Scheme 5.11: Preparation of ethyl thioglycosides.

The ethyl thioglycosides of mannose (**5.30**), galactose (**5.34**), GlcNAc (**5.38**), and lactose (**5.40**) were prepared using similar methodology to that followed previously (**Scheme 5.11**). Mannose ethyl thioglycoside **5.30** was prepared in an identical manner to glucose ethyl thioglycoside **5.22**, by first reacting α-mannose pentaacetate **5.28**[22] with boron trifluoride in

the presence of ethanethiol to afford protected thioglycoside **5.29**,[23] and then deprotecting under Zemplén conditions to afford *manno* ethyl thioglycoside **5.30**.[24] The *galacto* thioglycoside **5.34** was prepared similarly, however β-galactose pentaacetate **5.32** first had to be prepared, by acetylation of galactose **5.31** with acetic anhydride.[25] Subsequent conversion to thioglycoside **5.33**[26] and deprotection[21] was carried out as previously. The ethyl thioglycoside of GlcNAc **5.38** was prepared starting from the 2-phthalamido tetraacetate compound **5.35**, which was converted to the ethyl thioglycoside using boron trifluoride and ethanethiol as before to afford thioglycoside **5.36**.[23] The phthlamide group was then removed using hydrazine hydrate in ethanol and the resulting free amine acetylated with acetic anhydride to obtain fully acetylated thioglycoside **5.37**.[21] Zemplén deacetylation of **5.37** afforded unprotected GlcNAc ethyl thioglycoside **5.38**.[21] Lactose ethyl thioglycoside **5.40** was prepared from β-lactose octaacetate **5.39**[27] by treatment with boron trifluoride and ethanethiol as before.[28] The resulting protected thioglycoside could not be completely purified; a small amount of the α-thioglycoside was inseparable from the β anomer. Deprotection under Zemplén conditions and subsequent recrystallization afforded pure β-ethyl thioglycoside **5.40**.[28]

Pleasingly, in almost all cases high conversions were observed for the oxidation of the ethyl thioglycosides catalysed by flavinium salt **5.7** (**Scheme 5.12**, **Table 5.6**). The only exception was the GlcNAc ethyl thioglycoside **5.38**. NMR analysis of the crude reaction mixture showed both the presence of a significant amount of unreacted starting material and a number of products. As demonstrated previously protection of the thioglycoside significantly alters its reactivity in this reaction, and the poor performance of **5.38** would appear to indicate that protection or alteration of OH-2 in particular has a significant effect on the reactivity of the thioglycoside. This is perhaps unsurprising, as OH-2 is the hydroxyl group physically closest

to the anomeric centre. This is reinforced by the excellent reactivity of lactose thioglycoside **5.40**. Lactose thioglycoside **5.40** is insoluble in methanol, accounting for the low reactivity when methanol is used as a reaction solvent. Modification of a more remote hydroxyl group clearly does not have a significant effect on the reactivity of the thioglycoside substrate.



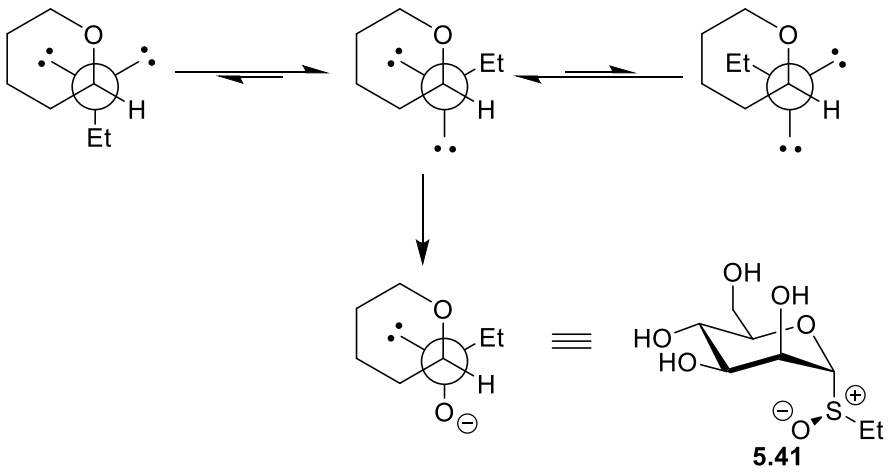
Scheme 5.12: Oxidation of ethyl thioglycosides to glycosyl sulfoxides.

Compound	Solvent	Conversion	Yield
5.22	MeOH	93%	76%
5.22	H_2O	>99%	71%
5.30	MeOH	93%	70%
5.30	H_2O	>99%	49%
5.34	MeOH	>99%	66%
5.34	H_2O	88%	63%
5.38	MeOH	Mixture	-
5.38	H_2O	Mixture	-
5.40	MeOH	-	-
5.40	H_2O	>99%	43%

Table 5.6: Oxidation of ethyl thioglycosides catalysed by **5.7**. In all cases a 5% loading of **5.7**, 1.2 eq. H_2O_2, and a reaction time of 6 h were used (**Scheme 5.12**).

The *manno* glycosyl sulfoxide **5.41** was unique in that it was obtained as a single diastereoisomer; all the other sulfoxides were formed as approximately 1:1 mixtures of the two sulfoxide diastereoisomers. The single diastereoisomer formed is believed to be the R_S diastereoisomer, as the preference for formation of the R_S diastereoisomer in oxidation reactions of α-thioglycosides has been noted previously.[29, 30] The explanation given for this phenomenon is that the favoured conformation of the thioglycoside α-anomer is one in which one lone pair of sulfur stabilises that conformation *via* the *exo*-anomeric effect, while the bulky thioglycoside substituent (the ethyl group in this case) is in the sterically favoured *exo* configuration. These two conditions lead to one favoured rotational conformer (**Scheme 5.13**). Oxidation then requires attack of one of the lone pairs on sulfur on the oxidising agent, in this case the flavinium-peroxide species. The pro-*S* lone pair is sterically hindered by the pyranose

ring, and so it is the pro-*R* lone pair that attacks preferentially, so forming a single diastereoisomer.[29] In addition, the conformer shown would also be expected to be the most reactive, as the perpendicular C-O and S-O dipoles would be expected to minimise the transition state energy. Thus, the conformation shown would be expected to be the most abundant and the most reactive.



5.41

Scheme 5.13: Mechanism of the selective formation of the R_S diastereoisomer of the *manno* ethyl glycosyl sulfoxide **5.41**.

As previously, the yields were in all cases lower than would be expected given the high conversions, and again this was due to difficulties in purifying the highly polar and relatively unstable glycosyl sulfoxide products. Beyond the difficulties of purification the relatively low stability of sulfoxides is rarely a factor in their use, as they are typically prepared and then used almost immediately. Nevertheless, once purified the glycosyl sulfoxides prepared here were found to be stable for weeks when stored at -20 °C.

5.8 Investigation of glycosyl sulfoxides as potential glycosidase inhibitors

5.8.1 Conditions

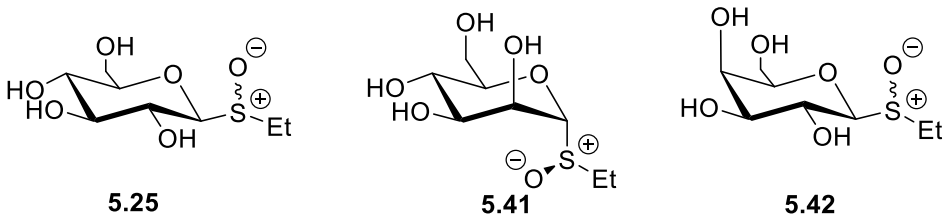


Figure 5.2: Sulfoxides selected for testing as glycosidase inhibitors.

A selection of the synthesised glycosyl sulfoxides were then tested as inhibitors of a variety of common glycosidase enzymes. The glycosyl sulfoxides selected were the *gluco, manno,* and *galacto* ethyl sulfoxides **5.25**, **5.41**, and **5.42** (**Figure 5.2**), which are potential inhibitors of β-glucosidases, α-mannosidases, and β-galactosidases respectively. The ethyl glycosyl sulfoxide series was the most easily prepared of the sulfoxides investigated, and they have a simple aglycon which would be expected to have minimal steric issues when the inhibitor is bound to an enzyme. It was thought that they therefore represented a good starting point from which modifications can be made if interesting inhibitory behaviour was discovered.

The glycosidase enzymes selected for these inhibition studies were a β-glucosidase from almonds,[31-33] a jack bean α-mannosidase,[34-36] and a β-galactosidase from *E. coli*.[37-39] All are well studied enzymes and are commercially available.

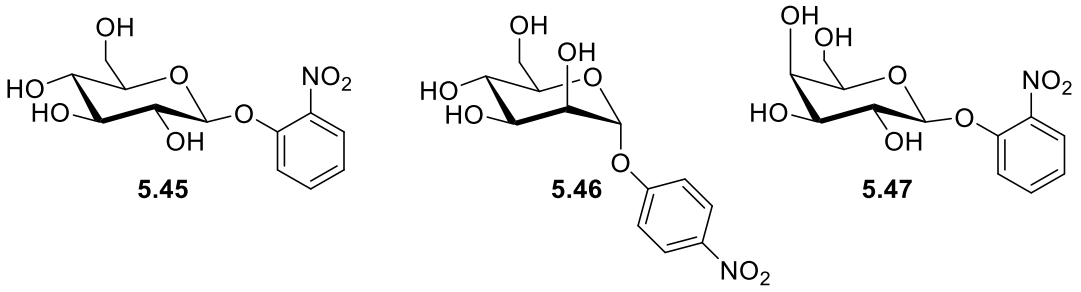


Figure 5.3: Nitrophenyl glycosides used for glycosidase inhibition studies.

Nitrophenyl glycosides **5.45**, **5.46**, and **5.47** (**Figure 5.3**) were chosen as hydrolytic substrates for the inhibition studies. Nitrophenyl glycosides are typically used for real time monitoring of the rate of the glycosidase catalysed hydrolysis reactions,[40, 41] as the nitrophenyl glycoside itself exhibits minimal or no absorption between 400 nm and 450 nm, while the *o*- or *p*-nitrophenolate ion produced by the hydrolysis reaction absorbs strongly in this region.

5.8.2 Kinetic model

One of two values is typically used to represent the strength of an enzyme inhibitor: either IC_{50}, the concentration of inhibitor required for the activity of the enzyme to be reduced by 50%, or K_i, the dissociation constant of the enzyme substrate complex.[42] The most significant difference between the two measurements is that K_i is by definition independent of substrate concentration, while IC_{50} is typically substrate concentration-dependent. Exceptions to the substrate-dependence of IC_{50} do exist, if the inhibitor is un- or non-competitive and certain conditions are met.[42] The substrate-independence of K_i means that K_i is normally a more useful measurement for comparing data between various studies. However, determination of K_i does require the collection of significantly more data than determination of IC_{50}, as initial rate values must be collected over a range of substrate and inhibitor concentrations.

A modified version of Michaelis-Menten enzyme kinetics is used to determine K_i.[43]

$$v = \frac{V_{max}[S]}{K_m + [S]} \quad (1)$$

Equation 1 is the standard Michaelis-Menten equation for an enzyme-catalysed reaction involving a single substrate. For a competitive inhibitor it can be assumed that K_m is modified by some factor that is dependent on K_i and the concentration of inhibitor, to give a lower apparent K_m value, K_m^{app} (Equation 2).

$$K_m^{app} = K_m \left(1 + \frac{[I]}{K_i}\right) \quad (2)$$

Substituting this value for K_m^{app} into equation 1 gives equation 3, which describes the initial rate in terms of V_{max}, [S], [I], K_m, and K_i.

$$v = \frac{V_{max}[S]}{\left(1 + \frac{[I]}{K_i}\right)K_m + [S]} \quad (3)$$

If measurements of the initial rate of reaction are taken at varying concentrations of both substrate and inhibitor Equation 3 can be fitted to the experimental data to give values for V_{max}, K_m, and K_i. The fitting of Equation 3 is best achieved by direct fitting to the data using the method of least squares.[44] It is possible to convert Equation 3 into a linear form analogous to the Lineweaver-Burk plot (equation 4).[45]

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{\left(1 + \frac{[I]}{K_i}\right)K_m}{V_{max}[S]} \quad (4)$$

The use of a linear plot results in a significantly higher error in fitted values, and therefore linear fitting to Equation 4 is not suitable for calculating K_i or K_m.[44] However, the use of a

linear plot does allow for more effective visual analysis of the data. In order to fit the data obtained in this study, the program GraFit 5 was used.

There are three common types of reversible enzyme inhibition.[46] The glycosyl sulfoxides examined here are expected to be competitive inhibitors, which means that substrate and inhibitor both bind to the enzyme active site, and therefore only one of the two can be bound at a time. Competitive inhibitors raise the apparent K_m of the enzyme, by reducing the concentration of enzyme available for reaction, but result in no change in $V_{max,}$, as the rate at which the available enzyme can process substrate remains unchanged. The type of inhibition is readily determined by observation of the Lineweaver-Burk plot. If the inhibition observed is competitive then the lines corresponding to various inhibitor concentrations will share a common y-intercept, corresponding to $\frac{1}{V_{max}}$.

In non-competitive inhibition the substrate and inhibitor bind to the enzyme in an independent manner, that is, both, one of, or neither of the substrate and inhibitor may be bound to the enzyme at one time. Non-competitive inhibition results in no change in apparent K_m, but a lowering of V_{max}, as substrate binding is not affected by the binding of inhibitor but the rate at which the enzyme can process substrate is reduced. In the case of non-competitive inhibition, when inhibitor concentration is varied the Lineweaver-Burk plot displays a series of lines with a common x-intercept, corresponding to the constant apparent K_m value. Mixed inhibition is related to non-competitive inhibition, and is when the strength of substrate binding to the uninhibited enzyme is different to the strength of substrate binding to the enzyme-inhibitor complex. When inhibitor concentration is varied the Lineweaver-Burk plot of the activity of

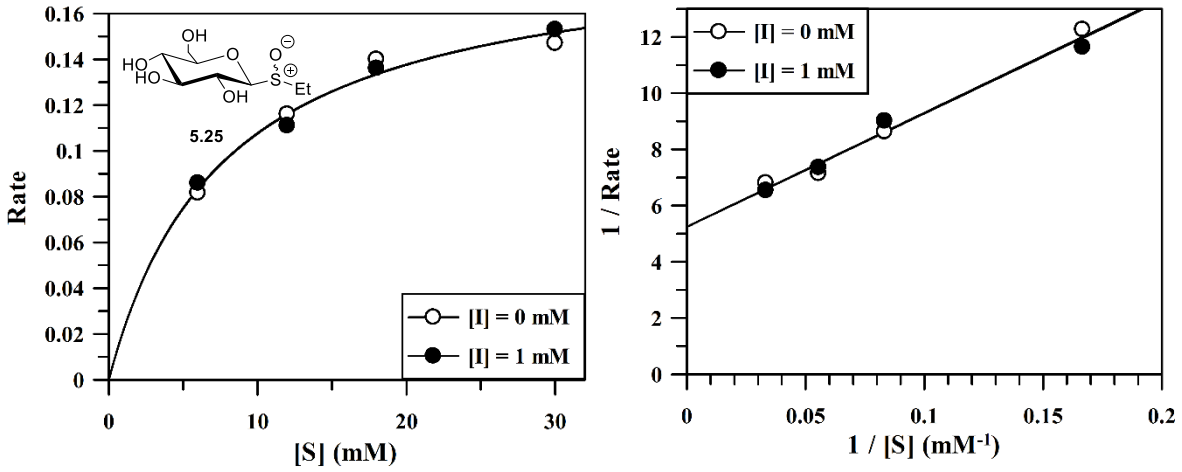
an enzyme inhibited by a mixed inhibitor displays a series of lines with a common intercept that is on neither the x- nor the y-axis.

The third common type of reversible inhibition is uncompetitive inhibition. An uncompetitive inhibitor only binds to the enzyme-substrate complex, not to the free enzyme. Uncompetitive inhibition results in a decrease in the rate at which the enzyme can process substrate and also results in a decrease in the apparent K_m of the enzyme, as the formation of enzyme-substrate-inhibitor complex effectively lowers the concentration of the enzyme substrate complex, increasing the amount of substrate which can bind to the enzyme. The Lineweaver-Burk plot of the activity of an uncompetitively inhibited enzyme displays a series of parallel lines, as both K_i and V_{max} vary when inhibitor concentration is varied.

While the glycosyl sulfoxides examined in this study were expected to be competitive inhibitors, uncompetitive inhibition of glycosidase enzymes by some glycosyl sulfoxides has been previously reported,[47] and so the possibility of inhibition by another mechanism was not discounted.

5.8.3 Inhibition testing

The first sulfoxide tested was the *gluco* sulfoxide **5.25**, which was assayed against β-glucosidase from almonds. The study was carried out at pH 6.8 in a phosphate buffer (0.05 M) at 25 °C.[41]

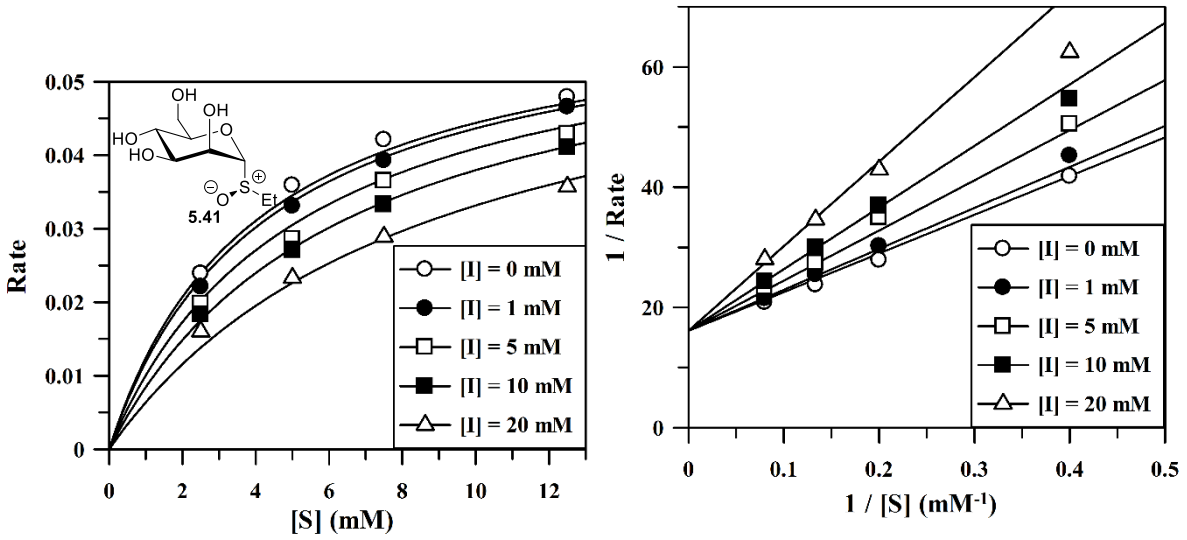


	Value	Error
V_{max}	0.191 s^{-1}	±0.08 s^{-1}
K_m	7.7 mM	±1.0 mM
K_i	-	-

Figure 5.4: Lack of inhibition of β-glucosidase from almonds by *gluco* sulfoxide **5.25**.

Unfortunately, **5.25** did not inhibit the β-glucosidase from almonds at an inhibitor concentration of 1 mM (**Figure 5.4**). The value of K_m obtained is broadly consistent with previous studies, which reported a K_m of 6.0 mM.[48]

Next, *manno* sulfoxide **5.41** was tested against the α-mannosidase from jack beans. The experiments were carried out at pH 4.5 in a sodium acetate buffer (0.1 M) at 37 °C.[40]



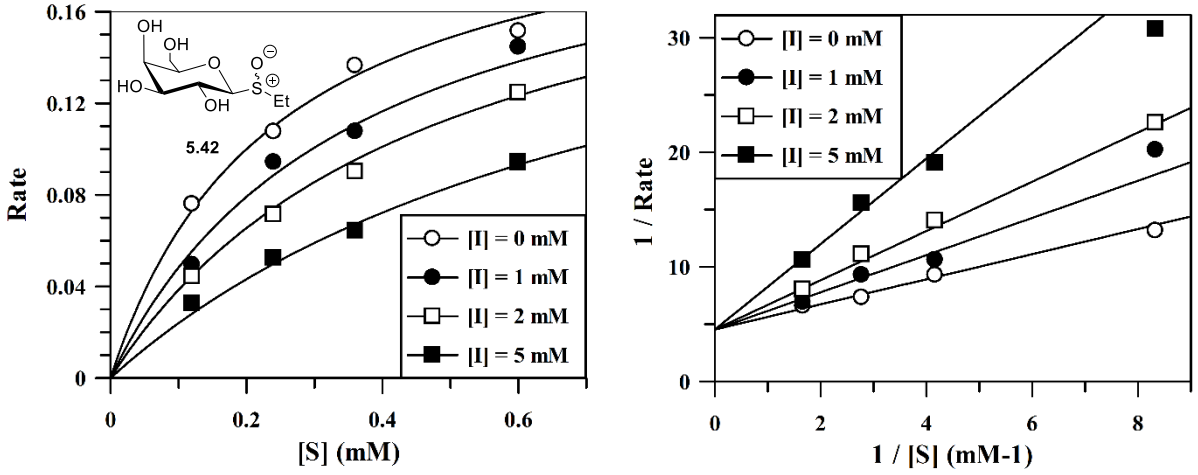
	Value	Error
V_{max}	0.062 s^{-1}	±0.002 s^{-1}
K_m	4.0 mM	±0.3 mM
K_i	17 mM	±1 mM

Figure 5.5: Inhibition of jack bean α-mannosidase by *manno* sulfoxide **5.41**.

The *manno* sulfoxide **5.41** gave slightly more promising results (**Figure 5.5**). Inhibition of jack bean α-mannosidase was observed at an inhibitor concentration of 1 mM, and further testing revealed a K_i for **5.41** of 17 mM. Again, the measured value of K_m of 4.0 mM was broadly consistent with the previously reported value of 2.5 mM.[49] Millimolar level inhibition of lysosomal and golgi α-mannosidases by α-mannosyl sulfoxides was previously reported by Poláková *et al.*,[50] results similar to those demonstrated here. In addition, the common y-

intercept at varying inhibitor concentrations on the Lineweaver-Burk plot indicated that the inhibition was competitive.

The *galacto* sulfoxide **5.42** was assayed against the β-galactosidase from *E. coli*, at pH 7.0 in a sodium phosphate buffer (0.1 M) at 37 °C.[51]



	Value	Error
V_{max}	0.22 s^{-1}	±0.01 s^{-1}
K_m	0.24 mM	±0.03 mM
K_i	2.1 mM	±0.02 mM

Figure 5.6: Inhibition of the β-galactosidase from *E. coli* by *galacto* sulfoxide **5.42**.

Pleasingly, *galacto* sulfoxide **5.42** also demonstrated mild inhibitory activity against the corresponding glycosidase, with a K_i of 2.1 mM obtained (**Figure 5.6**). The K_m value of 0.24 mM is consistent with previously reported values.[52] However, similarly to *manno* sulfoxide **5.41**, the *galacto* sulfoxide **5.42** displayed only moderate activity against the corresponding glycosidase. Again, the inhibition was competitive.

5.8.4 Inhibition testing of single sulfoxide diastereoisomers

One notable feature of the *gluco* and *galacto* sulfoxides tested is that they were mixtures of diastereoisomers, and, as enzyme active sites are chiral environments, it is likely that the inhibitory activities of the two diastereoisomers are in fact different. Therefore, the two diastereoisomers of each of the *gluco* and *galacto* sulfoxides were separated and the K_is of the individual diastereoisomers were measured.

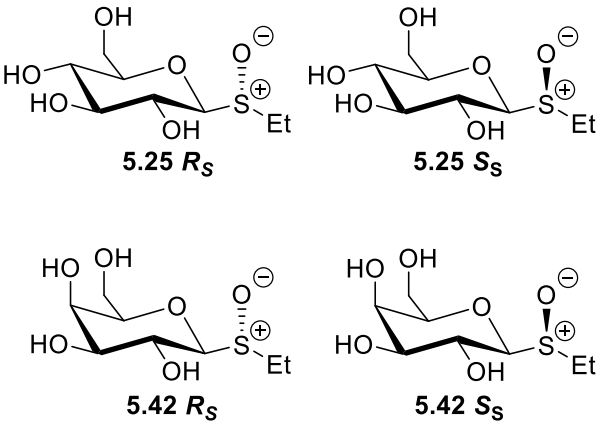


Figure 5.7: Sulfoxide diastereoisomers separated by HPLC.

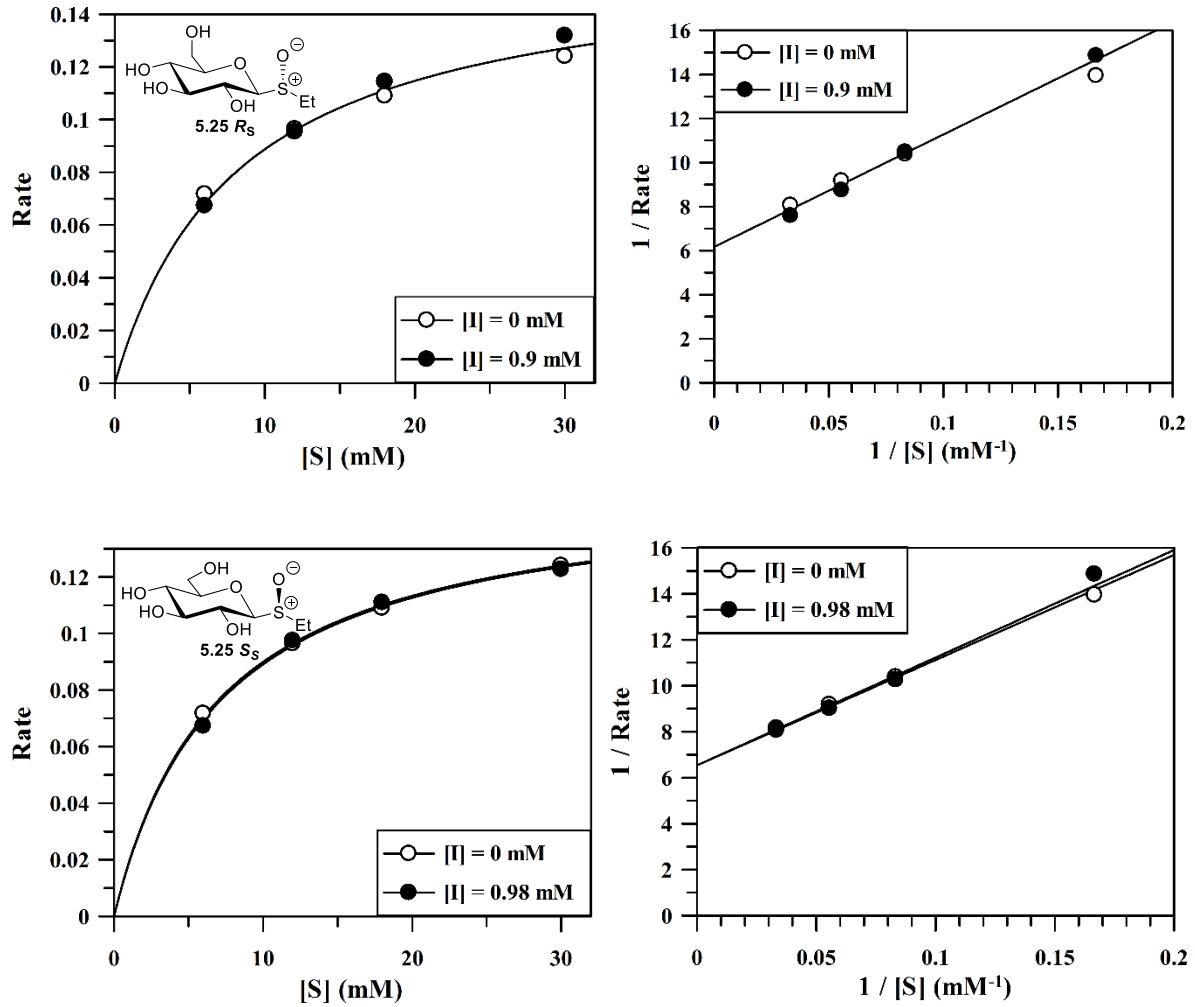
Both sets of diastereoisomers were separated by HPLC, as the sulfoxides are non-crystalline and were not separable by flash column chromatography, resulting in four new potential glycosidase inhibitors: *gluco* sulfoxides **5.25** ***S*S** and ***R*S**, and *galacto* sulfoxides **5.42** ***S*S** and ***R*S**. Determining which diastereoisomer was which was non-trivial, as they were not literature compounds, and so NMR spectra, optical rotation values, and other data were not available for comparison. There was also no significant preference for the formation of one diastereoisomer over the other, as was the case for the *manno* sulfoxide **5.41**.

Four techniques have been used to assign the configuration of glycosyl sulfoxides. X-Ray crystallography is the most reliable, and has often been used to verify the results of other

methods.[30, 53] Unfortunately, none of the sulfoxides obtained in this study produced large enough crystals for examination by this method. Chiral shift reagents have been used to distinguish sulfoxide diastereoisomers, although the difference between the two isomers in that case was less than 0.02 ppm for the single sulfoxide studied.[53] Circular dichroism has also been applied to a small group of sulfoxides in a study by Sanhueza *et al*.[54] While circular dichroism is effective a simpler method was preferable, ideally using NMR data already collected, as there are noticeable differences in the NMR spectra of the diastereoisomers.

Khair has reported that differences in chemical shift between the two diastereoisomers of ethyl glycosyl sulfoxides can reliably be used to assign the configuration of the sulfur.[55] Notably, the peak corresponding to C-1 of the R_S sulfoxide is shifted upfield relative to that of the S_S sulfoxide. In addition, the peak or peaks corresponding to the CH_2 group of the ethyl substituent display significantly higher separation in the R_S substituent as compared to the S_S substituent (roughly 160 Hz as compared to roughly 20 Hz). These effects are relatively consistent and are due to the varying degree to which conformers displaying the *exo*-anomeric effect are present in solution for the R_S and S_S diastereoisomers.[56] The R_S diastereoisomer favours only the conformation in which the *exo*-anomeric effect is present, resulting in greater shielding of C-1.

In the case of **5.25** and **5.42** the separation of the CH_2 peaks was not useful to distinguish the two diastereoisomers, as in all cases the separation between the peaks was approximately 50 Hz. However, the chemical shift C-1 varied in a manner consistent with the above method for distinguishing diastereoisomers, and so assignments were made on the basis of these chemical shifts.



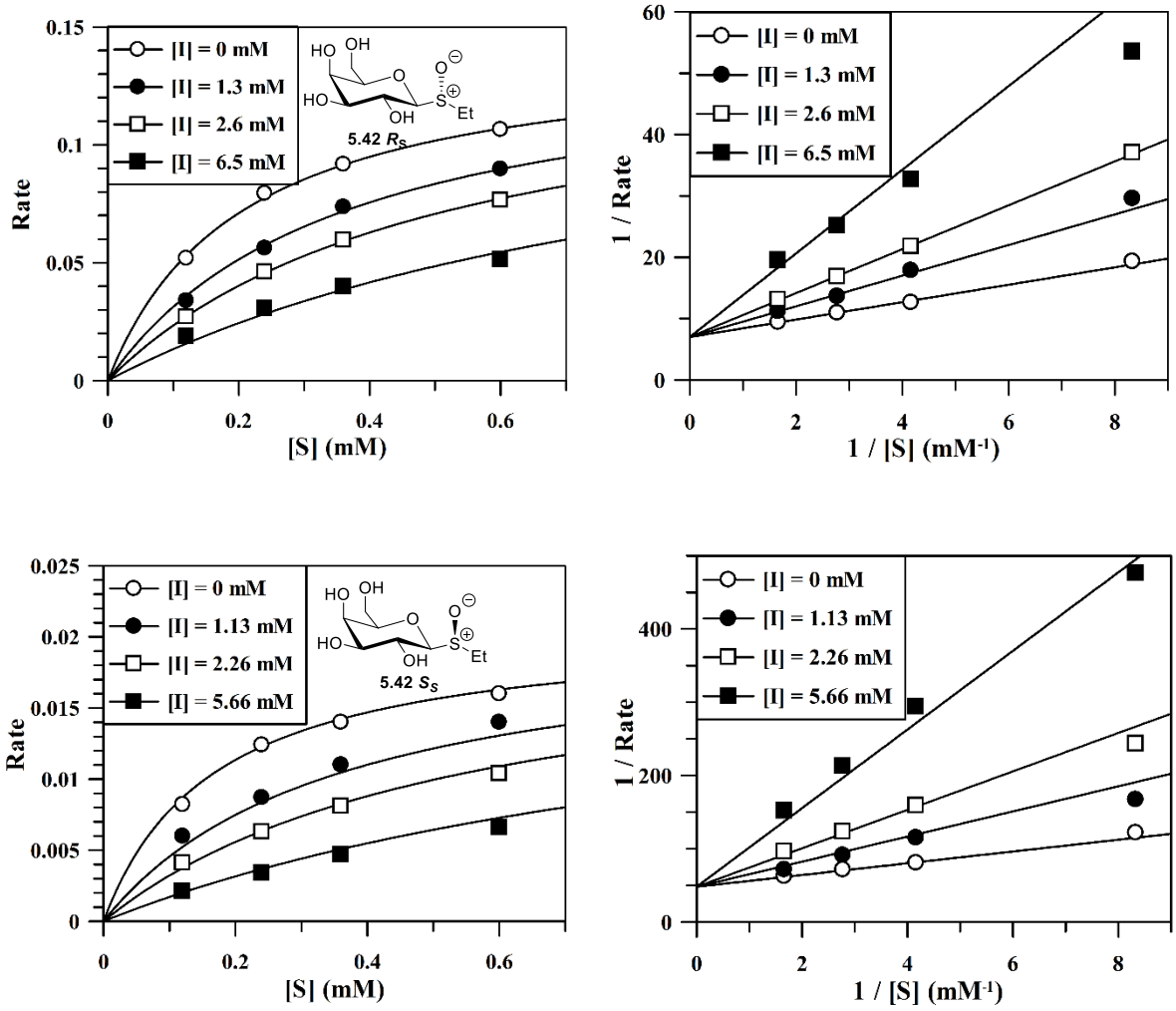
	5.25 R_S		5.25 S_S	
	Value	**Error**	**Value**	**Error**
V_{max}	0.162 s^{-1}	±0.006 s^{-1}	0.153 s^{-1}	±0.003 s^{-1}
K_m	8.3 mM	±0.9 mM	7.0 mM	±0.4 mM
K_i	-	-	-	-

Figure 5.8: Inhibition testing of *gluco* sulfoxides against the β-glucosidase from almonds. Glucosyl sulfoxides **5.25 *R*S** and ***S*S** were again assayed against β-glucosidase from almonds using the same conditions as previously. While the diastereoisomeric mixture did not potently inhibit the glucosidase, it is possible that one of the individual sulfoxides could display

significantly higher inhibitory activity than the other. Again, no statistically significant inhibition of β-glucosidase from almonds was observed at a 1 mM inhibitor concentration for either diastereoisomer (**Figure 5.8**).

Galacto sulfoxide **5.42** was previously shown to inhibit the β-galactosidase from *E. coli*. Testing of the individual diastereoisomers revealed that the inhibitory activity of sulfoxide **5.42** did indeed vary with the stereochemistry of the sulfoxide.

The results of the inhibition studies of the two diastereoisomers of **5.42** on the β-galactosidase from *E. coli* (**Figure 5.9**) indicated that both diastereoisomers inhibited the β-galactosidase, but that **5.42** $\boldsymbol{S_S}$ was more active than **5.42** $\boldsymbol{R_S}$, with K_is of 1.0 mM and 1.7 mM respectively. This observation is consistent with a previously reported study in which the S_S isomer of a sulfoxide-linked disaccharide was found to display greater inhibitory activity against the β-galactosidase from *E. coli* than the R_S isomer, with K_is of 0.19 mM and 0.45 mM respectively.[57] Again these values represent only moderate inhibition which, while it may have some mechanistic interest, is of limited use in the design of potent glycosidase inhibitors.



	5.42 *R*S		5.42 *S*S	
	Value	**Error**	**Value**	**Error**
V_{max}	0.143 s^{-1}	±0.004 s^{-1}	0.021 s^{-1}	±0.01 s^{-1}
K_m	0.20 mM	±0.02 mM	0.17 mM	±0.02 mM
K_i	1.7 mM	±0.1 mM	1.0 mM	±0.1 mM

Figure 5.9: Inhibition of β-galactosidase from *E. coli* by the *galacto* sulfoxide diastereoisomers.

5.9 Conclusions

In summary, a biomimetic oxidation process for the oxidation of unprotected thioglycosides to glycosyl sulfoxides was developed and applied to a range of thioglycosides. A selection of the glycosyl sulfoxides produced in this way were examined as inhibitors of glycosidase enzymes, where they displayed only moderate inhibitory activity.

A number of potential future avenues of exploration remain open. Firstly, only a very small subset of flavinium salt catalysts was examined in this study. Modification of the salt may provide a catalyst which is better able to catalyse the oxidation of protected thioglycosides in non-polar solvents. Secondly, the selection of an alternate catalyst may allow the use of oxygen as an oxidant, avoiding the requirement for potentially hazardous hydrogen peroxide. Finally, examination of a wider selection of glycosyl sulfoxides may potentially reveal more effective inhibitors of glycosidases. Notably, the ethyl aglycon is unlikely to interact favourably with any part of the enzyme active site, and selection of a more suitable aglycon could enhance the inhibitory activity of the sulfoxide significantly.

5.10 References

1. Sheldon, R. A., *Chem. Commun.* **2008,** (29), 3352-3365.

2. Murahashi, S. I.; Oda, T.; Masui, Y., *J. Am. Chem. Soc.* **1989,** *111* (13), 5002-5003.

3. Bergstad, K.; Bäckvall, J.-E., *J. Org. Chem.* **1998,** *63* (19), 6650-6655.

4. Marsh, B. J.; Carbery, D. R., *Tetrahedron Lett.* **2010,** *51* (17), 2362-2365.

5. Trost, B. M., *Angew. Chem., Int. Ed.* **1995,** *34* (3), 259-281.

6. Žurek, J.; Cibulka, R.; Dvořáková, H.; Svoboda, J., *Tetrahedron Lett.* **2010,** *51* (7), 1083-1086.

7. Li, W. S.; Zhang, N.; Sayre, L. M., *Tetrahedron* **2001,** *57* (21), 4507-4522.

8. Zemplen, G.; Kunz, A., *Ber. Dtsch. Chem. Ges.* **1923,** *56*, 1705-1710.

9. France, R. R.; Compton, R. G.; Davis, B. G.; Fairbanks, A. J.; Rees, N. V.; Wadhawan, J. D., *Org. Biomol. Chem.* **2004,** *2* (15), 2195-2202.

10. Aucagne, V.; Aversa, M. C.; Barattucci, A.; Bonaccorsi, P.; Giannetto, P.; Rollin, P.; Tatibouët, A., *J. Org. Chem.* **2002,** *67* (20), 6925-6930.

11. Liu, J.; Huang, C.-Y.; Wong, C.-H., *Tetrahedron Lett.* **2002,** *43* (18), 3447-3448.

12. Kakarla, R.; Dulina, R. G.; Hatzenbuhler, N. T.; Hui, Y. W.; Sofia, M. J., *J. Org. Chem.* **1996,** *61* (23), 8347-8349.

13. Minidis, A. B. E.; Bäckvall, J. E., *Chem. Eur. J.* **2001,** *7* (1), 297-302.

14. Ménová, P.; Cibulka, R., *J. Mol. Catal. A: Chem.* **2012,** *363–364*, 362-370.

15. Ménová, P.; Dvořáková, H.; Eigner, V.; Ludvík, J.; Cibulka, R., *Adv. Synth. Catal.* **2013,** *355* (17), 3451-3462.

16. Brenner, A., *Electrodeposition of Alloys: Principles and Practice*. Elsevier Science: 2013.

17. Ferrier, R. J.; Furneaux, R. H., 36 - 1,2-trans-1-Thioglycosides. In *General Methods*, BeMiller, R. L. W. N., Ed. Academic Press: 1980; pp 251-253.

18. Vibhute, A. M.; Muvvala, V.; Sureshan, K. M., *Angew. Chem., Int. Ed.* **2016,** *55* (27), 7782-7785.

19. Meng, B.; Zhu, Z.; Baker, D. C., *Org. Biomol. Chem.* **2014,** *12* (28), 5182-5191.

20. Li, X.; Huang, L.; Hu, X.; Huang, X., *Org. Biomol. Chem.* **2009,** *7* (1), 117-127.

21. Vic, G.; Hastings, J. J.; Howarth, O. W.; Crout, D. H. G., *Tetrahedron: Asymmetry* **1996,** *7* (3), 709-720.

22. Dowlut, M.; Hall, D. G.; Hindsgaul, O., *J. Org. Chem.* **2005,** *70* (24), 9809-9813.

23. Saibal Kumar, D.; Roy, N., *Carbohydr. Res.* **1996,** *296* (1), 275-277.

24. Wang, C.; Sanders, B.; Baker, D. C., *Can. J. Chem.* **2011,** *89* (8), 959-963.

25. Liu, Z.; Byun, H.-S.; Bittman, R., *Org. Lett.* **2010,** *12* (13), 2974-2977.

26. Mandal, S. S.; Liao, G.; Guo, Z., *RSC Advances* **2015,** *5* (30), 23311-23319.

27. Abellán Flos, M.; García Moreno, M. I.; Ortiz Mellet, C.; García Fernández, J. M.; Nierengarten, J.-F.; Vincent, S. P., *Chem. Eur. J.* **2016,** *22* (32), 11450-11460.

28. Tomoo, T.; Kondo, T.; Abe, H.; Tsukamoto, S.; Isobe, M.; Goto, T., *Carbohydr. Res.* **1996,** *284* (2), 207-222.

29. Crich, D.; Mataka, J.; Sun, S.; J. Wink, D.; Lam, K. C.; L. Rheingold, A., *Chem. Commun.* **1998,** (24), 2763-2764.

30. Crich, D.; Mataka, J.; Zakharov, L. N.; Rheingold, A. L.; Wink, D. J., *J. Am. Chem. Soc.* **2002,** *124* (21), 6028-6036.

31. Grover, A. K.; David MacMurchie, D.; Cushley, R. J., *Biochim. Biophys. Acta Enzymol.* **1977,** *482* (1), 98-108.

32. He, S.; Withers, S. G., *J. Biol. Chem.* **1997,** *272* (40), 24864-24867.

33. Weber, J. P.; Fink, A. L., *J. Biol. Chem.* **1980,** *255* (19), 9030-9032.

34. Snaith, S. M.; Levvy, G. A., *Biochem. J.* **1968,** *110* (4), 663-670.

35. Snaith, S. M., *Biochem. J.* **1975,** *147* (1), 83-90.

36. Hembre, E. J.; Pearson, W. H., *Tetrahedron* **1997,** *53* (32), 11021-11032.

37. Juers, D. H.; Heightman, T. D.; Vasella, A.; McCarter, J. D.; Mackenzie, L.; Withers, S. G.; Matthews, B. W., *Biochemistry* **2001,** *40* (49), 14781-14794.

38. Matthews, B. W., *C. R. Biol.* **2005,** *328* (6), 549-556.

39. Huber, R.; Kurz, G.; Wallenfels, K., *Biochemistry* **1976,** *15* (9), 1994-2001.

40. Tropea, J. E.; Molyneux, R. J.; Kaushal, G.; Pan, Y.; Mitchell, M.; Elbein, A. D., *Biochemistry* **1989,** *28* (5), 2027-2034.

41. Lohse, A.; Hardlei, T.; Jensen, A.; Plesner, I. W.; Bols, M., *Biochem. J.* **2000,** *349* (1), 211-215.

42. Yung-Chi, C.; Prusoff, W. H., *Biochem. Pharmacol.* **1973,** *22* (23), 3099-3108.

43. Berg, J. M.; Tymoczko, J. L.; Stryer, L., *Biochemistry, Fifth Edition*. W.H. Freeman: 2002.

44. Colquhoun, D., *J. R. Stat. Soc. Ser. C Appl. Stat.* **1969,** *18* (2), 130-140.

45. Lineweaver, H.; Burk, D., *J. Am. Chem. Soc.* **1934,** *56* (3), 658-666.

46. Copeland, R. A., *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*. 2nd ed.; John Wiley and Sons, Inc.: New York, 2000.

47. Robina, I.; Gomez-Bujedo, S.; Fernandez-Bolanos, J.; Del Pozo, L.; Demange, R.; Picasso, S.; Vogel, P., *Carbohydr. Lett.* **2000,** *3*, 389-396.

48. Barman, T. E., EC 3.2.1.21. In *Enzyme Handbook*, Berlin, 1969; Vol. 2, pp 578-579.

49. Li, Y.-T., *J. Biol. Chem.* **1967,** *242* (23), 5474-5480.

50. Poláková, M.; Šesták, S.; Lattová, E.; Petruš, L.; Mucha, J.; Tvaroška, I.; Kóňa, J., *Eur. J. Med. Chem.* **2011,** *46* (3), 944-952.

51. Lo, S.; Dugdale, M. L.; Jeerh, N.; Ku, T.; Roth, N. J.; Huber, R. E., *Protein J.* **2010,** *29* (1), 26-31.

52. Wallenfels, K.; Weil, R., 20 β-Galactosidase. In *The Enzymes*, Paul, D. B., Ed. Academic Press: 1972; Vol. Volume 7, pp 617-663.

53. Buist, P. H.; Behrouzian, B.; MacIsaac, K. D.; Cassel, S.; Rollin, P.; Imberty, A.; Gautier, C.; Pérez, S.; Genix, P., *Tetrahedron: Asymmetry* **1999,** *10* (15), 2881-2889.

54. Sanhueza, C. A.; Arias, A. C.; Dorta, R. L.; Vázquez, J. T., *Tetrahedron: Asymmetry* **2010,** *21* (15), 1830-1832.

55. Khiar, N., *Tetrahedron Lett.* **2000,** *41* (47), 9059-9063.

56. Sanhueza, C. A.; Dorta, R. L.; Vázquez, J. T., *J. Org. Chem.* **2011,** *76* (19), 7769-7780.

57. Colomer, J. P.; Mayordomo, M. Á. C.; De Toro, B. F.; Jiménez-Barbero, J.; Varela, O., *Eur. J. Org. Chem.* **2015,** *2015* (7), 1448-1455.

Chapter 6: Experimental

General experimental section

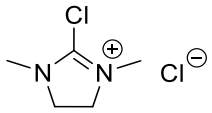
All reactions involving moisture sensitive reagents were performed under an atmosphere of argon or nitrogen *via* standard vacuum Schlenk line techniques. All glassware for such reactions was flame-dried and cooled under an atmosphere of argon. Reactions conducted at -20 °C were cooled by means of an acetonitrile/dry ice bath; those conducted at 0 °C were cooled by means of an ice bath. Solvent was removed under reduced pressure using a Buchi™ rotary evaporator. HPLC-grade solvents were used for reactions, and in case of moisture-sensitive reactions solvents were dried by literature procedures and freshly distilled as required. Petroleum ether (Petrol) refers to the fraction of light petroleum ether boiling in the range 40-60 °C. Reagents were used as supplied without further purification unless otherwise stated. Thin Layer Chromatography (t.l.c.) was carried out on Merck Silica Gel 60F254 aluminium-backed plates. Visualisation of the plates was achieved using a UV lamp (λ_{max} = 254 or 365 nm), and/or ammonium molybdate (5% in 2 M aqueous H_2SO_4) and/or potassium permanganate (60 mM in aqueous 0.5 M potassium carbonate and 2.5 mM sodium hydroxide) and /or aniline-diphenylamine-85% phosphoric acid (4 mL: 4 g: 20 mL) in acetone (96 mL). Flash column chromatography was carried out using Sorbsil C60 40/60 silica. Melting points were recorded on an Electrothermal melting point apparatus and are uncorrected. Proton and carbon nuclear magnetic resonance (^{1}H, ^{13}C) spectra were recorded on Agilent 400–MR instrument operating for ^{1}H NMR at 400 MHz, and at 100.5 MHz for ^{13}C NMR, or an Agilent 600–MR instrument operating for ^{1}H NMR at 600 MHz, and at 150 MHz for ^{13}C NMR. All chemical shifts are quoted on the δ-scale in ppm using residual solvent as an internal standard. ^{1}H and ^{13}C spectra were assigned using COSY, DEPT, HSQC, HMBC, and TOCSY. High resolution mass spectra were recorded by on either a DIONEX Ultimate 3000 or Bruker MaXis

4G spectrometer, operated in high resolution positive ion electrospray mode. Optical rotations were measured on a Perkin-Elmer 241 polarimeter with a 1 cm^3 cell with a path length of 1 dm, and are quoted in units of °.cm^2.g^{-1}. Concentrations (c) are given in g/100 cm^3, solvent and temperature are recorded. Infrared spectra were recorded on a Bruker FTIR spectrometer with Alpha's Platinum ATR single reflection diamond where the neat samples were recorded.

Lysozyme from hens' egg white, BSA, β-glucosidase from almonds, β-galactosidase from *Eschericia coli*, and α-mannosidase from jack beans were purchased from Sigma-Aldrich and used as received. *N*-butoxycarbonyl-L-cysteine-L-threonine methyl ester **3.22** was prepared by Patrycja Gołębiowska. GABARAPL2 was prepared by Dr. Akshita Wason. Peptide **3.39** was prepared by Zaid Anso. Sialoglycan **3.41** was prepared by Dr. Vivek Poonthiyil. Modified MUC1 peptide **4.30** was prepared by Dr. Geoff Williams. Endo M N175Q was purchased from Tokyo Chemical Industry and used as recieved. *p*-Tolyl-2,3,4,6-*O*-acetyl-1-thio-β-D-glucopyranoside **5.8**, 1,3,4,6-tetra-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside **5.35**, and 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-1,2,3,6-tetra-*O*-acetyl-β-D-glucopyranoside **5.39** were prepared by previous members of the Fairbanks group.

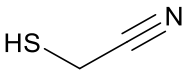
Experimental section for Chapter 2

2-Chloro-1,3-dimethylimidazolinium chloride 2.1[1]



1,3-Dimethylimidazolidinone (50 mL, 475 mmol) was dissolved in toluene (500 mL) and stirred at 50 °C under an N_2 atmosphere. Oxalyl chloride (51 mL, 594 mmol) was added dropwise. After 3 h, the flask was allowed to cool to RT and stirred overnight. After 18 h, the precipitate was filtered, washed with cold toluene and recrystallized ($MeCN/Et_2O$). The white solid was filtered and dried under vacuum to yield 2-chloro-1,3-dimethylimidazolidinium chloride **2.1** (37.27 g, 45 %) as a white powdery solid; m.p. 65-70 °C [lit. 95-100 °C][1]; δ_H (400 MHz, CD_3CN) 3.15 (6H, s, 2 x CH_3), 3.98 (4H, s, 2 x CH_2); δ_C (100.5 MHz, CD_3CN) 34.3, 49.9; HRMS (ESI-TOF): calcd. for $C_5H_{10}ClN_2^+$: 133.0533. Found: 133.0528 (M^+).

Mercaptoacetonitrile 2.2



Method 1[2]

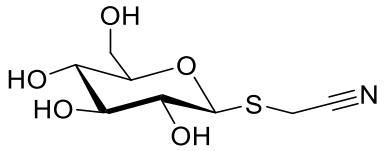
A solution of sodium hydrosulfide hydrate (25.3 g, 0.27 mol) in water (43 mL) was cooled to -20 °C and stirred under an atmosphere of nitrogen. Chloroacetonitrile (21.6 mL, 0.24 mol) was added rapidly and the reaction was stirred at -20 °C for 5 h, at which point NaCl was observed to precipitate from the reaction mixture. The reaction was then warmed to RT over 1 h and H_2SO_4 (conc., 3 mL) was added to bring the reaction mixture to pH 1. The solution was then diluted with water (20 mL), and extracted with CH_2Cl_2 (3 x 100 mL). The combined organic extracts were concentrated *in vacuo* and the crude product was purified by flash column chromatography (petrol:ethyl acetate, 5:1) to afford mercaptoacetonitrile **2.2** (3.62 g, 21%) as

a yellow liquid; υ_{max} (neat) 2247 cm^{-1} (C≡N); δ_H (400 MHz, $CDCl_3$)[2] 2.37 (1H, t, *J* 8.2 Hz, SH), 3.29 (2H, d, *J* 8.2 Hz, CH_2); δ_C (100.5 MHz, $CDCl_3$) 9.6 (t, $\underline{C}H_2$), 118.2 (s, $\underline{C}N$).

Method 2[3]

Chloroacetonitrile (6.3 mL, 0.1 mol) and thioacetic acid (7.05 mL, 0.1 mol) were dissolved in CH_2Cl_2 (70 mL) and cooled to -30 °C under an atmosphere of N_2. Triethylamine (16.7 mL, 0.12 mol) was added dropwise over 0.5 h. The reaction was stirred at -30 °C for another 1 h, and then gradually warmed to RT. Water (5 mL) was added, and the organic layer was washed with acetic acid (10%, 2 x 15 mL) followed by water (2 x 15 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The liquid produced was dissolved in MeOH (100 mL) and Amberlyst® 15 (H^+ form, 3 g) was added. The reaction was heated to reflux under an atmosphere of N_2 and stirred for 20 h. The reaction was filtered, Amberlyst® 15 (H^+ form, 0.3 g) was added, and the mixture was concentrated *in vacuo* at RT to afford mercaptoacetonitrile **2.2** (5.96 g, 82%) as a pale yellow liquid, identical in all respects to that produced by *Method 1*.

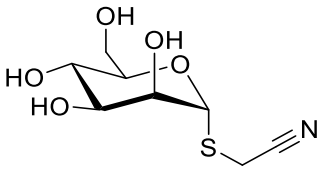
Cyanomethyl 1-thio-β-D-glucopyranoside 2.4



D-Glucose **2.3** (200 mg, 1.11 mmol) was dissolved in water (4 mL). Triethylamine (2.5 mL, 34 mmol) was added, the reaction was cooled to 0 °C, and stirred. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (0.92 g, 5.5 mmol) was dissolved in mercaptoacetonitrile **2.2** (1.3 g, 17.8 mmol) and the resulting solution was added dropwise to the aqueous D-glucose solution over 2 min. The reaction was stirred at 0 °C for 0.5 h, and was then diluted with water

(5 mL) and washed with CH_2Cl_2 (5 x 10 mL). The aqueous layer was filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), lyophilised, pre-absorbed onto silica, and purified by flash column chromatography ($CHCl_3$:MeOH, 5:1) to afford cyanomethyl 1-thio-β-D-glucopyranoside **2.4** (232 mg, 89%) as a pale yellow oil; $[\alpha]_D^{20}$ -67 (*c*, 1.0 in MeOH) [lit. $[\alpha]_D^{22}$ -60.2 (*c*, 5.05 in water)][4] ; υ_{max} (neat) 2248 cm^{-1} (C≡N); δ_H (400 MHz, D_2O) 3.26 - 3.35 (2H, m, H-2 & H-3), 3.37 - 3.44 (2H, m, H-4 & H-5), 3.57, 3.68 (2H, ABq, *J* 17.6 Hz, C$\underline{H}_2$CN) 3.61 (1H, dd, $J_{5,6}$ 5.5 Hz, $J_{6,6'}$ 12.1 Hz, H-6), 3.79 (1H, dd, $J_{5,6'}$ 1.8 Hz, $J_{6,6'}$ 12.3 Hz, H-6'), 4.60 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O) 14.4 (t, $\underline{C}H_2CN$), 60.7 (t, C-6), 69.2 (d, C-2), 71.8 (d, C-3), 77.0, 80.0 (2 x d, C-4 & C-5), 84.4 (d, C-1), 118.6 (s, $CH_2\underline{C}N$); HRMS (ESI-TOF): calcd. for $C_8H_{13}NO_5SNa^+$: 258.0407. Found: 258.0414 (MNa^+).

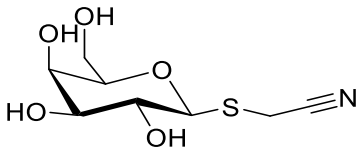
Cyanomethyl 1-thio-α-D-mannopyranoside 2.5



D-Mannose (100 mg, 0.56 mmol) was dissolved in water (2 mL). Triethylamine (1.2 mL, 11.1 mmol) was added, the reaction was cooled to 0 °C, and stirred. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (0.56 g, 3.3 mmol) was dissolved in mercaptoacetonitrile **2.2** (0.8 g, 11.1 mmol) and the resulting solution was added dropwise to the aqueous D-mannose solution over 2 min. The reaction was stirred at 0 °C for 0.5 h, and was then diluted with water (2 mL) and washed with CH_2Cl_2 (5 x 10 mL). The aqueous layer was filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), lyophilised, pre-absorbed onto silica, and purified by flash column chromatography ($CHCl_3$:MeOH, 5:1) to afford cyanomethyl 1-thio-

α-D-mannopyranoside **2.5** (103 mg, 78%) as a yellow oil; $[\alpha]_D^{20}$ +80 (*c*, 1.0 in MeOH); υ_{max} (neat) 3327 cm^{-1} (OH), 2255 cm^{-1} (CN); δ_H (400 MHz, D_2O) 3.49, 3.54 (2H, ABq, *J* 17.6 Hz, SC<u>H</u>$_2$CN), 3.71 - 3.76 (2H, m, H-3 & H-4), 3.80 (2H, dd, $J_{5,6}$ 5.5 Hz, $J_{6,6'}$ 12.5 Hz, H-6), 3.88 (2H, dd, $J_{5,6'}$ 2.0 Hz, $J_{6,6'}$ 12.1 Hz, H-6'), 3.91 - 3.97 (1H, m, H-5), 4.08 (1H, d, $J_{2,3}$ 1.2 Hz, H-2), 5.47 (1H, s, H-1); δ_C (100.5 MHz, D_2O) 15.3 (t, S<u>C</u>H$_2$CN), 60.4 (t, C-6), 66.7 (d, C-3), 70.8 (d, C-2), 71.0 (d, C-4), 73.5 (d, C-5), 84.8 (d, C-1), 118.2 (s, SCH$_2$<u>C</u>N); HRMS (ESI-TOF): calcd. for $C_8H_{13}NO_5SNa^+$: 258.0407. Found: 258.0401 (MNa^+).

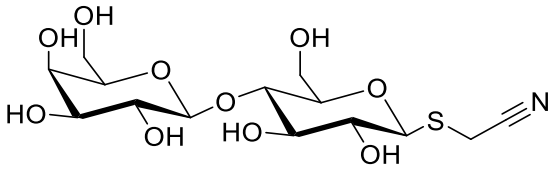
Cyanomethyl 1-thio-β-D-galactopyranoside 2.6



D-Galactose (200 mg, 1.11 mmol) was dissolved in water (4 mL). Triethylamine (2.4 mL, 33 mmol) was added, the reaction was cooled to 0 °C, and stirred. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (1.12 g, 6.7 mmol) was dissolved in mercaptoacetonitrile **2.2** (1.6 g, 22 mmol) and the resulting solution was added dropwise to the aqueous D-galactose solution over 2 min. The reaction was stirred at 0 °C for 0.5 h, and was then diluted with water (5 mL) and washed with CH_2Cl_2 (5 x 10 mL). The aqueous layer was filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), lyophilised, pre-absorbed onto silica, and purified by flash column chromatography ($CHCl_3$:MeOH, 5:1) to afford cyanomethyl 1-thio-β-D-galactopyranoside **2.6** (106 mg, 41%) as a pale yellow oil; $[\alpha]_D^{20}$ +13 (*c*, 0.1 in MeOH) [lit. $[\alpha]_D^{20}$ -51.5 (*c*, 5.05 in water)][4]; υ_{max} (neat) 2252 cm^{-1} (C≡N); δ_H (400 MHz, D_2O) 3.50 - 3.68 (5H, m, H-2, H-4, H-5, H-6 & H-6'), 3.57, 3.69 (2H, ABq, *J* 17.6 Hz, CH_2) 3.87 (1H, d, $J_{2,3}$ 3.1 Hz, H-3), 4.54 (1H, d, $J_{1,2}$ 9.4 Hz, H-1); δ_C (100.5 MHz, D_2O) 14.5 (t, <u>C</u>H$_2$CN), 61.0

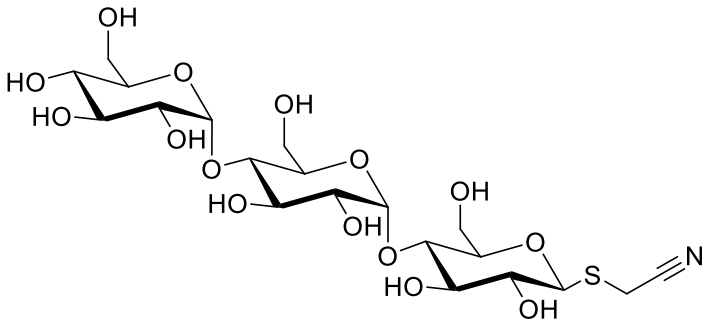
(t, H-6), 68.6 (d, C-3), 69.2 (d, C-2), 73.8, 79.1 (2 x d, C-4 & C-5), 84.9 (d, C-1), 118.6 (s, $CH_2\underline{C}N$); HRMS (ESI-TOF): calcd. for $C_8H_{13}NO_5SNa^+$: 258.0407. Found: 258.0413 (MNa^+).

Cyanomethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 2.7



β-D-Galactopyranosyl-(1→4)-D-glucopyranose (103 mg, 0.3 mmol) was dissolved in water (2 mL). Triethylamine (0.84 mL, 6 mmol) was added, the reaction was cooled to 0 °C, and stirred. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (0.303 g, 1.8 mmol) was dissolved in mercaptoacetonitrile **2.2** (0.439 g, 6 mmol) and the resulting solution was added dropwise to the aqueous β-D-galactopyranosyl-(1→4)-D-glucopyranose solution over 2 min. The reaction was stirred at 0 °C for 0.5 h, and was then diluted with water (2 mL) and washed with CH_2Cl_2 (5 x 10 mL). The aqueous layer was filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), lyophilised, pre-absorbed onto silica, and purified by flash column chromatography ($CHCl_3$:MeOH, 5:1) to afford cyanomethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.7** (77 mg, 65%) as a yellow oil; $[\alpha]_D^{20}$ -25 (*c*, 0.1 in MeOH); υ_{max} (neat) 3298 cm^{-1} (OH), 2251 cm^{-1} (CN); δ_H (400 MHz, CD_3OD) 3.37 (1H, t, *J* 9.8 Hz, H-2_a), 3.42 - 3.65 (6H, m, H-3_a, H-4_a, H-5_a, H-2_b, H-3_b & H-5_b), 3.66 - 3.88 (6H, m, $C\underline{H}_2$, H-6_b, H-4_b, H-6_b & H-6'$_b$), 3.94 (1H, dd, $J_{5,6}$ 2.0 Hz, $J_{6,6'}$ 12.1 Hz, H-6'$_a$), 4.37 (1H, d, $J_{1,2}$ 7.4 Hz, H-1_b), 4.60 (1H, d, $J_{1,2}$ 9.4 Hz, H-1_a); δ_C (100.5 MHz, CD_3OD) 13.1 (t, $\underline{C}H_2CN$), 60.5 (t, C-6_a), 61.1 (t, C-6_b), 68.9 (d, C-4_b), 71.1, 72.6, 73.3, 75.6, 76.4, 78.8, 79.4 (7 x d, C-2_a, C-3_a, C-4_a, C-5_a, C-2_b, C-3_b & C-5_b), 83.9 (d, C-1_a), 103.6 (d, C-1_b), 117.3 (s, $CH_2\underline{C}N$); HRMS (ESI-TOF): calcd. for $C_{14}H_{24}NO_{10}S^+$: 398.1115. Found: 398.1109 (MH^+).

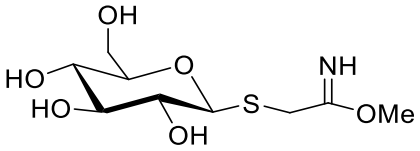
Cyanomethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 2.8



α-D-Glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-β-D-glucopyranose (101 mg, 0.2 mmol) was dissolved in D_2O (1.3 mL). Triethylamine (0.56 mL, 4 mmol) was added, the reaction was cooled to 0 °C, and stirred. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (0.202 g, 1.2 mmol) was dissolved in mercaptoacetonitrile **2.2** (0.292 g, 4 mmol) and the resulting solution was added dropwise to the aqueous α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-β-D-glucopyranose solution over 2 min. The reaction was stirred at 0 °C for 0.5 h, and was then diluted with water (2 mL) and washed with CH_2Cl_2 (5 x 10 mL). The aqueous layer was filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), lyophilised, pre-absorbed onto silica, and purified by flash column chromatography ($CHCl_3$:MeOH, 5:1) to afford cyanomethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.8** (77 mg, 65%) as a yellow oil; $[\alpha]_D^{20}$ +54 (*c*, 1.0 in MeOH); υ_{max} (neat) 3279 cm^{-1} (OH), 2248 cm^{-1} (CN); δ_H (600 MHz, CD_3OD) 3.26 (1H, t, *J* 9.2 Hz, H-4_b), 3.33 (1H, t, *J* 9.6 Hz, H-2_a), 3.42 - 3.47 (2H, m, H-5_a & H-2_b), 3.48 - 3.53 (2H, m, H-2_c & H-5_c), 3.57 (1H, t, *J* 9.2 Hz, H-4_a), 3.61 (1H, t, *J* 9.4 Hz, H-3_b), 3.63 - 3.70 (3H, m, H-3_a, H-5_b & H-6_b), 3.73 - 3.84 (6H, m, H-6_a, H-6'$_b$, H-4_c, H-6_c & C$\underline{H}_2$), 3.84 - 3.88 (2H, m, H-3_c & H-6'$_c$), 3.90 (1H, dd, *J* 12.4 Hz, *J* 2.0 Hz, H-6'$_a$), 4.59 (1H, d, $J_{1,2}$ 9.6 Hz, H-1_a), 5.15 (1H, d, $J_{1,2}$ 3.7 Hz, H-1_b), 5.19 (1H, d, $J_{1,2}$ 3.9 Hz, H-1_c); δ_C (151 MHz, CD_3OD) 13.0 (t,

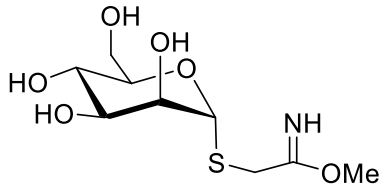
S<u>C</u>H$_2$CN), 60.7 (t, C-6$_c$), 60.9 (t, C-6$_a$), 61.3 (t, C-6$_b$), 70.1 (d, C-4$_b$), 71.9 (d, C-4$_c$), 72.4 (d, C-5$_c$), 72.5 (d, C-2$_a$), 72.8 (d, C-5$_b$), 73.4 (d, C-3$_c$), 73.5 (d, C-3$_b$), 73.6 (d, C-5$_a$), 77.8 (d, C-3$_a$), 79.3 (d, C-4$_a$), 79.4 (d, C-2$_b$), 79.9 (d, C-2$_c$), 83.9 (d, C-1$_a$), 101.2 (d, C-1$_c$), 101.5 (d, C-1$_b$), 117.2 (SCH$_2$<u>C</u>N); HRMS (ESI-TOF): calcd. for $C_{20}H_{33}NO_{15}SNa^+$: 582.14633. Found: 582.1461 (MNa^+).

2-Imino-2-methoxyethyl 1-thio-β-D-glucopyranoside 2.11[4]



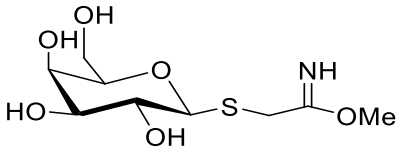
Sodium metal (8 mg, 0.36 mmol) was dissolved in MeOH (3.6 mL). Cyanomethyl 1-thio-β-D-glucopyranoside **2.4** (83.6 mg, 0.36 mmol) was added and the reaction mixture was stirred for 64 h. The reaction mixture was then concentrated *in vacuo* and analysed by ^{1}H NMR to reveal the formation of 2-imino-2-methoxyethyl 1-thio-β-D-glucopyranoside **2.11** (35% conversion). The crude material was subsequently used without purification for conjugation experiments; υ_{max} (neat) 1654 cm^{-1} (C=N); δ_H (400 MHz, CD_3OD) 3.17 - 3.41 (4H, m, H-2, H-3, H-4 & H-5)), 3.63 (1H, dd, $J_{5,6}$ 5.5 Hz, $J_{6,6'}$ 12.9 Hz, H-6), 3.85 (1H, d, $J_{6,6'}$ 12.9 Hz, H-6') 4.32 Hz (1H, d, *J* 10.2 Hz, H-1) ; δ_C (100.5 MHz, CD_3OD) 61.5 (t, C-6), 70.1 (d, C-2), 73.0 (d, C-3), 78.3, 80.9 (2 x d, C-4 & C-5), 85.2 (d, C-1); HRMS (ESI-TOF): calcd. for $C_9H_{18}NO_6S^+$: 268.0849. Found: 268.0860 (MH^+).

2-Imino-2-methoxyethyl 1-thio-α-D-mannopyranoside 2.12[4]



Sodium metal (13.7 mg, 0.60 mmol) was dissolved in MeOH (6 mL). Cyanomethyl 1-thio-α-D-mannopyranoside **2.5** (140 mg, 0.60 mmol) was added and the reaction mixture was stirred for 64 h. The reaction mixture was then concentrated *in vacuo* and analysed by ^{1}H NMR to reveal the formation of 2-imino-2-methoxyethyl 1-thio-α-D-mannopyranoside **2.12** (22% conversion). The crude material was subsequently used without purification for conjugation experiments; υ_{max} (neat) 3311 cm^{-1} (OH), 1649 cm^{-1} (C=NH); δ_H (400 MHz, CD_3OD) 3.59 - 3.84 (5H, H-3, H-4, H-6 & C$\underline{H}_2$), 3.87 (1H, d, $J_{6,6'}$ 13.7 Hz, H-6'), 3.91 (2H, m, H-2 & H-5), 5.33 (1H, s, H-1); δ_C (100.5 MHz, D_2O) 61.3 (t, C-6), 66.9 (d, C-3), 70.1 (d, C-2), 71.0 (d, C-4), 73.8 (d, C-5), 85.7 (d, C-1); HRMS (ESI-TOF): calcd. for $C_9H_{18}NO_6S^+$: 268.0849. Found: 268.0848 (MH^+).

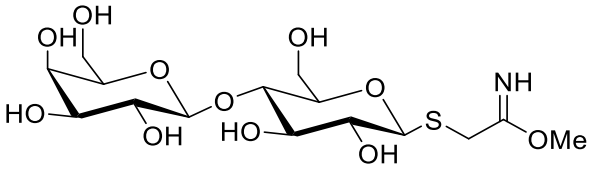
2-Imino-2-methoxyethyl 1-thio-β-D-galactopyranoside 2.13[4]



Sodium metal (5 mg, 0.22 mmol) was dissolved in MeOH (2.5 mL). Cyanomethyl 1-thio-β-D-galactopyranoside **2.6** (52 mg, 0.22 mmol) was added and the reaction mixture was stirred for 64 h. The reaction mixture was then concentrated *in vacuo* and analysed by ^{1}H NMR to reveal the formation of 2-imino-2-methoxyethyl 1-thio-β-D-galactopyranoside **2.13** (50% conversion). The crude material was subsequently used without purification for the conjugation experiments; υ_{max} (neat) 1651 cm^{-1} (C=N); δ_H (400 MHz, CD_3OD)[5] 3.40 - 3.61 (3H, m, H-2,

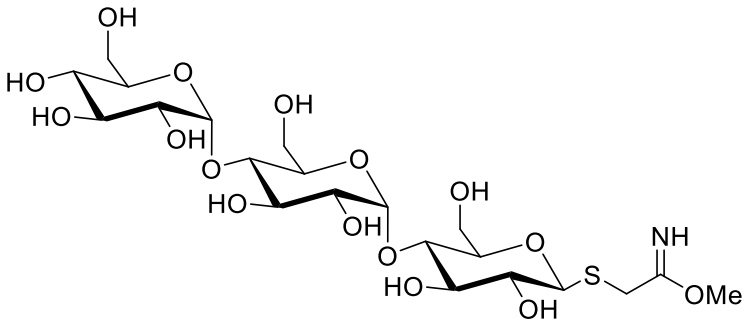
H-4 & H-5), 3.64 - 3.77 (3H, m, C<u>H</u>$_2$, H-6 & H-6'), 3.86 (1H, d, $J_{2,3}$ 10.6 Hz, $J_{3,4}$ 3.1 Hz, H-3) 4.27 (1H, d, $J_{1,2}$ 9.4 Hz, H-1); δ_C (100.5 MHz, CD_3OD) 61.3 (t, C-6), 69.2 (d, C-3), 70.0 (d, C-2), 74.9, 79.4 (2 x d, C-4 & C-5), 85.8 (d, C-1); HRMS (ESI-TOF): calcd. for $C_9H_{18}NO_6S^+$: 268.0849. Found: 268.0853 (MH^+).

2-Imino-2-methoxyethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 2.14[4]



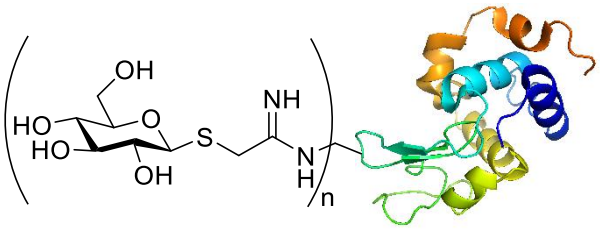
Sodium metal (7.6 mg, 0.33 mmol) was dissolved in MeOH (3.3 mL). Cyanomethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.7** (130 mg, 0.33 mmol) was added and the reaction mixture was stirred for 64 h. The reaction mixture was then concentrated *in vacuo* and analysed by ^{1}H NMR to reveal the formation of 2-imino-2-methoxyethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.14** (85% conversion). The crude material was subsequently used without purification for conjugation experiments; υ_{max} (neat) 3279 cm^{-1} (OH), 1652 cm^{-1} (C=NH); δ_H (400 MHz, CD_3OD) 4.35 (1H, d, *J* 7.4 Hz, H-1$_b$), 4.42 (1H, d, *J* 7.4 Hz, H-1$_a$); δ_C (100.5 MHz, CD_3OD) 60.5 (t, C-6$_a$), 60.5 (t, <u>C</u>H$_2$), 61.2 (t, C-6$_b$), 69.2 (d, C-4$_b$), 71.1, 72.2, 73.6, 75.8, 76.5, 78.7, 79.3 (7 x d, C-2$_a$, C-3$_a$, C-4$_a$, C-5$_a$, C-2$_b$, C-3$_b$ & C-5$_b$), 85.1 (d, C-1$_a$), 103.8 (d, C-1$_b$)) HRMS (ESI-TOF): calcd. for $C_{15}H_{28}NO_{11}S^+$: 430.1378. Found: 430.1392 (MH^+).

2-Imino-2-methoxyethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 2.15



Sodium metal (4 mg, 0.17 mmol) was dissolved in MeOH (1.7 mL). cyanomethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.8** (97 mg, 0.17 mmol) was added and the reaction mixture was stirred for 64 h. The reaction mixture was then concentrated *in vacuo* and analysed by ^{1}H NMR to reveal the formation of 2-imino-2-methoxyethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.15** (60% conversion). The crude material was subsequently used without purification for conjugation experiments; υ_{max} (neat) 3289 cm^{-1} (OH), 1652 cm^{-1} (C=NH); δ_H (400 MHz, D_2O) 4.69 (1H, d, $J_{1,2}$ 10.6 Hz, H-1$_a$), 5.29 - 5.35 (2H, m, H-1$_b$ & H-1$_c$); HRMS (ESI-TOF): calcd. for $C_{21}H_{38}NO_{16}S^+$: 592.1906. Found: 592.1907 (MH^+).

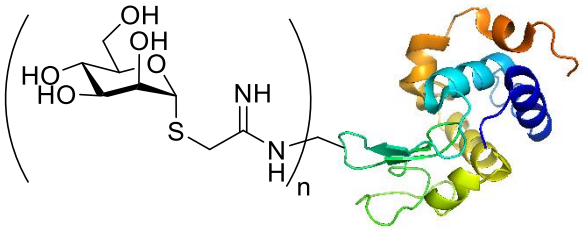
Conjugation of 2-imino-2-methoxyethyl 1-thio-β-D-glucopyranoside 2.11 to lysozyme



Lysozyme (1 mg, 0.070 µmol) was dissolved in sodium borate buffer (0.2 mL, 0.2 M, pH 8.5). A portion of the 2-imino-2-methoxyethyl 1-thio-β-D-glucopyranoside **2.11** crude reaction mixture (50 µL, approximately 1.8 µmol of **2.11**) was added to the lysozyme solution and the reaction was stirred. After 5 h the reaction was dialysed against water (2 x 500 mL), and

filtered; HRMS (ESI-TOF): 15246, 54% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_4$: 15247), 15481, 100% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_5$: 15482), 15715, 54% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_6$: 15717).

Conjugation of 2-imino-2-methoxyethyl 1-thio-α-D-mannopyranoside 2.12 to lysozyme



Method 1

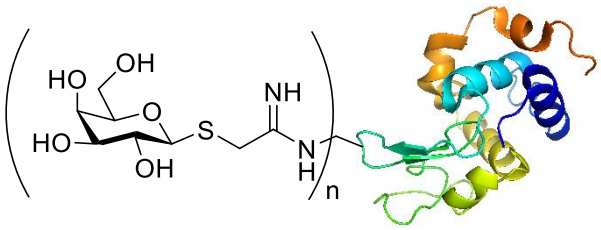
Lysozyme (1 mg, 0.070 µmol) was dissolved in sodium borate buffer (0.2 mL, 0.2 M, pH 8.5). A portion of the 2-imino-2-methoxyethyl 1-thio-α-D-mannopyranoside **62** crude reaction mixture (50 µL, approximately 1.1 µmol of **62**) was added to the lysozyme solution and the reaction was stirred. After 5 h the reaction was dialysed against water (2 x 500 mL), and filtered; HRMS (ESI-TOF): 14305, 25% (calcd. for lysozyme: 14307), 14540, 87% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)$: 14542), 14775, 100% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_2$: 14777), 15011, 47% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_3$: 15012), 15247, 6% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_4$: 15247).

Method 2

Lysozyme (1 mg, 0.070 µmol) was dissolved in sodium borate buffer (0.2 mL, 0.2 M, pH 8.5). A portion of the 2-imino-2-methoxyethyl 1-thio-α-D-mannopyranoside **62** crude reaction mixture (100 µL, approximately 2.2 µmol of **62**) was added to the lysozyme solution and the reaction was stirred. After 5 h the reaction was dialysed against water (2 x 500 mL), and filtered; HRMS (ESI-TOF): 14775, 2% (calcd. for lysozyme-$(C(NH)CH_2SC_6H_{11}O_5)_2$: 14777)

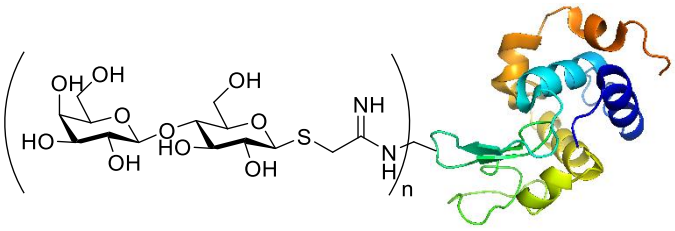
15011, 78% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_3$: 15013), 15249, 100% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_4$: 15247), 15482, 40% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_5$: 15482), 15717, 2% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_6$: 15717).

Conjugation of 2-imino-2-methoxyethyl 1-thio-β-D-galactopyranoside 2.13 to lysozyme



Lysozyme (1 mg, 0.070 µmol) was dissolved in sodium borate buffer (0.2 mL, 0.2 M, pH 8.5). A portion of the 2-imino-2-methoxyethyl 1-thio-β-D-galactopyranoside **2.13** crude reaction mixture (50 µL, approximately 2.2 µmol of **2.13**) was added to the lysozyme solution and the reaction was stirred. After 5 h the reaction was dialysed against water (2 x 500 mL), and filtered; HRMS (ESI-TOF): 15010, 2% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_3$: 15012), 15246, 61% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_4$: 15247), 15481, 100% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_5$: 15482), 15717, 54% (calcd. for lysozyme-(C(NH)$CH_2SC_6H_{11}O_5$)$_6$: 15717).

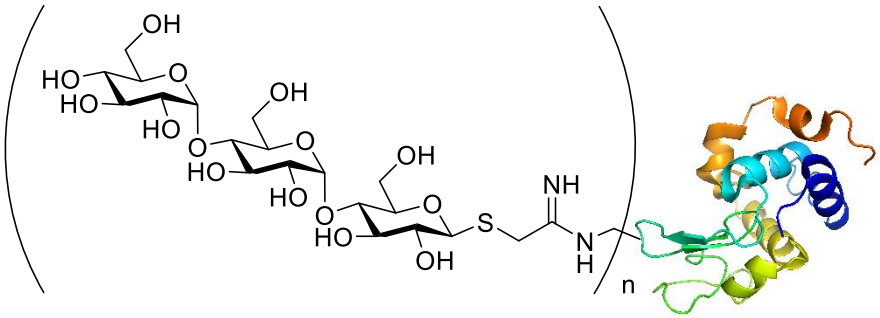
Conjugation of 2-imino-2-methoxyethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 2.14 to lysozyme



Lysozyme (1 mg, 0.070 µmol) was dissolved in sodium borate buffer (0.2 mL, 0.2 M, pH 8.5). A portion of the 2-imino-2-methoxyethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-

glucopyranoside **2.14** crude reaction mixture (50 µL, approximately 1.7 µmol of **2.14**) was added to the lysozyme solution and the reaction was stirred. After 5 h the reaction was dialysed against water (2 x 500 mL), and filtered; HRMS (ESI-TOF): 14307, 9% (calcd. for lysozyme: 14307), 14703, 25% (calcd. for lysozyme-($C(NH)CH_2SC_{12}H_{21}O_{10}$): 14704), 15099, 78% (calcd. for lysozyme-($C(NH)CH_2SC_{12}H_{21}O_{10}$)$_2$: 15101), 15497, 100% (calcd. for lysozyme-($C(NH)CH_2SC_{12}H_{21}O_{10}$)$_3$: 15498), 15894, 47% (calcd. for lysozyme-($C(NH)CH_2SC_{12}H_{21}O_{10}$)$_4$: 15895), 16293, 27% (calcd. for lysozyme-($C(NH)CH_2SC_{12}H_{21}O_{10}$)$_5$: 16292).

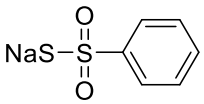
Conjugation of 2-imino-2-methoxyethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 2.15 to lysozyme



Lysozyme (1 mg, 0.070 µmol) was dissolved in sodium borate buffer (0.2 mL, 0.2 M, pH 8.5). A portion of crude 2-imino-2-methoxyethyl α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **2.15** (3 mg, approximately 3 µmol of **2.15**) was added to the lysozyme solution and the reaction was stirred. After 5 h the reaction was dialysed against water (2 x 500 mL), and filtered; HRMS (ESI-TOF): 14305, 9% (calcd. for lysozyme: 14307), 14864, 61% (calcd. for lysozyme-($C(NH)CH_2SC_{18}H_{31}O_{15}$): 14866), 150424, 100% (calcd. for lysozyme-($C(NH)CH_2SC_{18}H_{31}O_{15}$)$_2$: 15425), 15983, 50% (calcd. for lysozyme-($C(NH)CH_2SC_{18}H_{31}O_{15}$)$_3$: 15984), 16543, 15% (calcd. for lysozyme-($C(NH)CH_2SC_{18}H_{31}O_{15}$)$_4$: 16543), 17103, 1% (calcd. for lysozyme-($C(NH)CH_2SC_{18}H_{31}O_{15}$)$_5$: 17102).

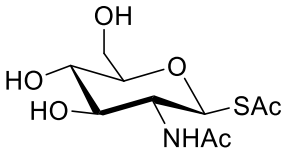
Experimental section for Chapter 3

Sodium phenylthiosulfonate 3.1[6]



Sodium benzenesulfinate (8.21 g, 50 mmol) and elemental sulfur (1.60 g, 50 mmol) were dissolved in pyridine (50 mL) under an atmosphere of nitrogen. The reaction was stirred at room temperature for 1 h, then the mixture was filtered and the precipitate was washed with Et_2O (50 mL). The precipitate was recrystallised from EtOH to afford sodium phenylthiosulfonate **3.1** (7.67 g, 78%) as a white crystalline solid; m.p. 295-296 °C (EtOH) [lit. 305-306 °C (EtOH)][6]; δ_H (400 MHz, DMSO-d_6)[6] 7.28 - 7.39 (3H, m, 2 x *m*-H & *p*-H), 7.75 (2H, d, *J* 6.7 Hz, 2 x *o*-H).

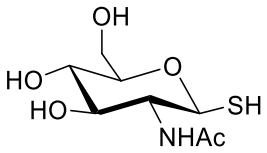
1-Thioacetyl-2-acetamido-2-deoxy-β-D-glucopyranoside 3.4



N-Acetyl-D-glucosamine **3.2** (500 mg, 2.26 mmol) and triethylamine (3.1 mL, 22.6 mmol) were stirred in water (10 mL) and cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (1.2 g, 6.78 mmol) was added. After 0.5 h, t.l.c. ($CHCl_3$:MeOH, 2:1) indicated complete consumption of starting material (R_f 0.2) and the formation of a major product (R_f 0.6). Thioacetic acid (2.4 mL, 33.9 mmol) was added dropwise and the reaction mixture was allowed to stir for an additional 0.5 h at 0 °C. The reaction mixture was then diluted with water (10 mL) and washed with CH_2Cl_2 (5 x 20 mL). The aqueous layer was concentrated *in vacuo* and the resulting solid dissolved in the minimum amount of water and filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous

NaOH solution and then washing with water until pH 7). The solution was concentrated *in vacuo*, dissolved in the minimum amount of water, and purified by flash column chromatography ($CHCl_3$:MeOH 5:1) to afford 1-thioacetyl-acetamido-2-deoxy-β-D-glucopyranoside **3.4** (0.361 g, 57%) as a white solid; m.p. 179-181 °C ($CHCl_3$); $[\alpha]_D^{20}$ +27 (*c*, 1.0 in MeOH); υ_{max} (neat) 3237 cm^{-1} (OH), 1688 cm^{-1} (SC=O), 1649 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.89 (3H, m, NHC(O)C$\underline{H}_3$), 2.31 (3H, s, SC(O)C$\underline{H}_3$), 3.39 (1H, t, *J* 9.0 Hz, H-4), 3.45 - 3.49 (1H, m, H-5), 3.52 (1H, t, *J* 9.0 Hz, H-3), 3.62 (1H, dd, $J_{5,6}$ 5.5 Hz, $J_{6,6'}$ 12.5 Hz, H-6), 3.74 - 3.83 (2H, m, H-2 & H-6'), 5.10 (1H, d, $J_{1,2}$ 10.6 Hz, H-1); δ_C (100.5 MHz, D_2O) 22.0 (q, NHC(O)$\underline{C}H_3$), 30.2 (q, SC(O)$\underline{C}H_3$), 53.4 (d, C-2), 60.5 (t, C-6), 69.4 (d, C-4), 74.9 (d, C-3), 80.3 (d, C-5), 80.9 (d, C-1), 174.5 (s, NH$\underline{C}$(O)CH_3), 197.6 (s, S$\underline{C}$(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{10}H_{18}NO_6S^+$: 280.0849. Found: 280.0848 (MH^+).

1-Thio-2-acetamido-2-deoxy-β-D-glucopyranose 3.5



Method 1

1-Thioacetyl-acetamido-2-deoxy-β-D-glucopyranoside **3.4** (279 mg, 1 mmol) was dissolved in aqueous NaOH (1 M, 3 mL) and stirred for 1 h. Amberlite® IR-120 (H^+) was then added until pH 7 was reached. The solution was then filtered and concentrated *in vacuo* to afford 1-thio-2-acetamido-2-deoxy-β-D-glucopyranose **3.5** (236 mg, quant.) as a white solid; m.p. 174-175 °C (ethyl acetate/MeOH) [lit. 177-179 °C (ethyl acetate/MeOH)][7]; $[\alpha]_D^{20}$ -6 (*c*, 1.0 in MeOH) [lit. $[\alpha]_D^{22}$ -10.4 (*c*, 1.0 in MeOH)][7]; υ_{max} (neat) 3286 cm^{-1} (OH, NH), 2531 cm^{-1} (SH), 1634 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.96 (3H, s, NHC(O)C$\underline{H}_3$), 3.35 - 3.48 (3H, m, H-3, H-4 & H-5), 3.59 - 3.70 (2H, m, H-2 & H-6), 3.78 - 3.84 (1H, m, H-6'), 4.60 (1H, d, $J_{1,2}$ 9.8 Hz,

H-1); δ_C (100.5 MHz, D_2O) 22.2 (q, NHC(O)<u>C</u>H_3), 58.0 (d, C-2), 60.8 (t, C-6), 69.7 (d, C-4), 75.0 (d, C-3), 79.1 (d, C-1), 80.1 (d, C-5), 174.6 (s, NH<u>C</u>(O)CH_3); HRMS (ESI-TOF): calcd. for $C_8H_{16}NO_5SNa^+$: 260.0563. Found: 260.0560 (MNa^+).

Method 2

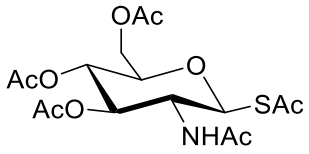
N-Acetyl-D-glucosamine **3.2** (500 mg, 2.26 mmol) and triethylamine (3.1 mL, 22.6 mmol) were stirred in water (10 mL) and the mixture cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (1.2 g, 6.78 mmol) was added. After 0.5 h, thioacetic acid (2.4 mL, 33.9 mmol) was added dropwise, and the reaction mixture was allowed to stir for an additional 0.5 h at 0 °C. The reaction mixture was then diluted with water (10 mL) and washed with CH_2Cl_2 (5 x 20 mL). The aqueous layer was concentrated *in vacuo* and the residue triturated with ethyl acetate (3 x 50 mL). The solid was then filtered off and dissolved in aqueous NaOH (1 M, 30 mL). The solution was stirred at RT for 1 h and then concentrated *in vacuo*. The residue was dissolved in water (50 mL) and Amberlite® IR-120 (H^+) was added until pH 7 was reached. The mixture was then filtered, and the filtrate was concentrated *in vacuo*. EtOH (50 mL) was added to the residue and the mixture was stirred. The mixture was filtered and the filtrate was concentrated *in vacuo* to afford 1-thio-2-acetamido-2-deoxy-β-D-glucopyranose **3.5** (0.453 g, 85%) as a white solid, identical in all respects to that produced by *Method 1*.

Method 3

Sodium metal (25.4 mg, 1.1 mmol) was added to MeOH (5 mL) and stirred. 1-Thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-β-D-glucopyranoside **3.6** (203 mg, 0.50 mmol) was added and the reaction mixture was stirred for 0.5 h, at which point t.l.c. (ethyl acetate:petrol 2:1) indicated complete consumption of starting material (R_f 0.2) and formation of a single

product (R_f 0). Dowex® 50WX8 (H^+) ion exchange resin was added portion-wise until the reaction reached neutral pH. The mixture was then filtered and concentrated *in vacuo* to afford 1-thio-2-acetamido-2-deoxy-β-D-glucopyranose **3.5** (122 mg, quant.) as a white solid, identical in all respects to that produced by *method 1*.

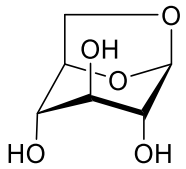
1-Thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-β-D-glucopyranoside 3.6



N-Acetyl-D-glucosamine **3.2** (500 mg, 2.26 mmol) and triethylamine (3.1 mL, 22.6 mmol) were stirred in water (10 mL), and cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (1.2 g, 6.78 mmol) was added. After 0.5 h, thioacetic acid (2.4 mL, 33.9 mmol) was added dropwise, and the reaction mixture was allowed to stir for an additional 0.5 h at 0 °C. The reaction mixture was then diluted with water (10 mL) and washed with DCM (5 x 20 mL). The aqueous layer was concentrated *in vacuo* and the solid suspended in pyridine (7 mL) under an atmosphere of nitrogen. The mixture was stirred at RT and acetic anhydride (7 mL, 74.1 mmol) was added. The reaction was stirred for 12 h, after which time t.l.c. (petrol:ethyl acetate 2:1) indicated the formation of a single major product (R_f 0.8). Water (10 mL) was added, and the mixture was then washed with CH_2Cl_2 (3 x 100 mL). The combined organic extracts were washed with aqueous HCl (1 M, 100 mL), $NaHCO_3$ (sat. aqueous soln., 100 mL), water (100 mL) and brine (100 mL), dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The residue was recrystallised (CH_2Cl_2/petrol) to afford 1-thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-β-D-glucopyranoside **3.6** (0.807 g, 88%) as a colourless crystalline solid; m.p. 190-191 °C (CH_2Cl_2/petrol) [lit. 186-188 °C (CH_2Cl_2/petrol)][8]; $[\alpha]_D^{20}$ +20 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ +12.1 (*c*, 2.2 in $CHCl_3$)][9]; υ_{max} (neat) 1739 cm^{-1} (OC=O), 1699 cm^{-1} (SC=O), 1661 cm^{-1} (NHC=O); δ_H (400 MHz, $CDCl_3$) 1.92 (3H, s, NHC(O)C<u>H</u>$_3$), 2.03 (6H, s, 2 x C(O)C<u>H</u>$_3$), 2.08

(3H, s, C(O)C<u>H</u>$_3$), 2.37 (3H, s, SC(O)C<u>H</u>$_3$), 3.75 - 3.81 (1H, m, H-5), 4.09 (1H, dd, $J_{6,6'}$ 11.7 Hz, H-6), 4.23 (1H, dd, $J_{6,6'}$ 12.7 Hz, $J_{5,6'}$ 4.5 Hz, H-6'), 4.35 (1H, aq, J 9.9 Hz, H-2), 5.05 - 5.18 (3H, m, H-1, H-3 & H-4), 5.65 (1H, d, $J_{NH,2}$ 9.8 Hz, N<u>H</u>C(O)CH$_3$); δ_C (100.5 MHz, $CDCl_3$) 20.6 (q, C(O)<u>C</u>H$_3$), 20.6 (q, C(O)<u>C</u>H$_3$), 20.7 (q, C(O)<u>C</u>H$_3$), 23.1 (q, NHC(O)<u>C</u>H$_3$), 30.8 (q, SC(O)<u>C</u>H$_3$), 52.3 (d, C-2), 61.8 (t, C-6), 67.7 (d, C-3), 74.0 (d, C-4), 76.6 (d, C-5), 81.6 (d, C-1), 169.2 (s, <u>C</u>(O)CH$_3$), 170.0 (s, <u>C</u>(O)CH$_3$), 170.7 (s, <u>C</u>(O)CH$_3$), 171.3 (s, <u>C</u>(O)CH$_3$), 193.7 (s, S<u>C</u>(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{16}H_{24}NO_9S^+$: 406.1172. Found: 406.1166 (MH^+).

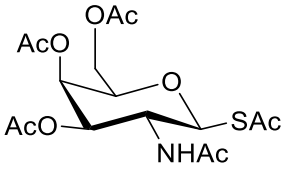
1,6-Anhydro-β-D-glucose 3.9[10]



Glucose **3.7** (180 mg, 1.0 mmol) and triethylamine (1.37 mL, 10 mmol) were stirred in water (4.4 mL) and cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (504 mg, 3.0 mmol) was added. After 0.5 h, thioacetic acid (1.06 mL, 15 mmol) was added dropwise and the reaction mixture was stirred for an additional 0.5 h at 0 °C. The reaction mixture was then diluted with water (5 mL) and washed with CH_2Cl_2 (5 x 10 mL). The aqueous layer was filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7). The solution was concentrated *in vacuo*, dissolved in the minimum amount of water, and purified by flash column chromatography ($CHCl_3$:MeOH 5:1) to afford 1,6-anhydro-β-D-glucose **3.9** (65 mg, 40%) as a colourless oil; δ_H (400 MHz, D_2O)[10] 3.50 (1H, s, H-2), 3.63 - 3.68 (2H, m, H-3 & H-4), 3.73 (1H, dd, $J_{5,6}$ 5.9 Hz, $J_{6,6'}$ 7.8 Hz, H-6), 4.07 (1H, d, $J_{6,6'}$ 7.8 Hz, H-6'), 4.61 (1H, d, $J_{5,6}$ 5.9 Hz, H-5), 5.43 (1H, s, H-1); δ_C (100.5 MHz, D_2O) 65.1 (t, C-6), 70.1 (d, C-2), 70.8 (d,

C-4), 72.4 (d, C-3), 76.2 (d, C-5), 101.4 (d, C-1); HRMS (ESI-TOF): calcd. for $C_6H_{10}O_5Na^+$: 185.0420. Found: 185.0417 (MNa^+).

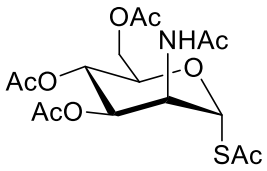
1-Thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-β-D-galactopyranoside 3.13



N-Acetyl-D-galactosamine **3.10** (500 mg, 2.26 mmol) and triethylamine (3.1 mL, 22.6 mmol) were stirred in water (10 mL) and the mixture was cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (1.2 g, 6.78 mmol) was added. After 0.5 h, thioacetic acid (2.4 mL, 33.9 mmol) was added dropwise and the reaction mixture was allowed to stir at 0 °C for an additional 0.5 h. The reaction mixture was then diluted with water (10 mL) and washed with CH_2Cl_2 (5 x 20 mL). The aqueous layer was concentrated *in vacuo* and the solid stirred in pyridine (7 mL) under an atmosphere of nitrogen. Acetic anhydride (7 mL, 74.1 mmol) was added and the mixture was stirred for 12 h, after which time t.l.c. (ethyl acetate:petrol, 2:1) indicated the formation of a single product (R_f 0.1). Water (10 mL) was added, and the mixture was then washed with CH_2Cl_2 (3 x 100 mL). The combined organic extracts were washed with aqueous HCl (1 M, 100 mL), $NaHCO_3$ (sat. aqueous soln., 100 mL), water (100 mL) and brine (100 mL), dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (ethyl acetate:petrol, 2:1) to afford 1-thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-β-D-galactopyranoside **3.13** (511 mg, 56%) as a colourless solid; m.p. 185-187 °C (CH_2Cl_2/Et_2O) [lit. 197-198 °C (CH_2Cl_2/Et_2O)];[11] $[\alpha]_D^{20}$ +20 (*c*, 1.0 in $CHCl_3$) [lit. +11.3 (*c*, 1.0 in $CHCl_3$)][11]; υ_{max} (neat) 1738 cm^{-1} (OC=O), 1701 cm^{-1} (SC=O), 1645 cm^{-1} (NHC=O); δ_H (400 MHz, $CDCl_3$)[11] 1.92 (3H, s, C(O)C$\underline{H}_3$), 2.01 (3H, s, C(O)C$\underline{H}_3$), 2.04 (3H, s, C(O)C$\underline{H}_3$), 2.16 (3H, s, C(O)C$\underline{H}_3$), 2.38 (3H, s, SC(O)C$\underline{H}_3$), 3.98 - 4.16 (3H, m, H-5, H-6 & H-6'), 4.52 (1H, q, *J* 10.6 Hz, H-2), 5.05 (1H, dd, $J_{2,3}$ 10.8 Hz,

$J_{3,4}$ 3.3 Hz, H-3), 5.15 (1H, d, $J_{1,2}$ 10.6 Hz, H-1), 5.36 - 5.45 (2H, m, NHC(O)CH_3 & H-4); δ_C (100.5 MHz, $CDCl_3$) 20.6 (q, C(O)CH_3), 20.7 (q, C(O)CH_3), 20.7 (q, C(O)CH_3), 23.2 (q, C(O)CH_3), 30.8 (q, SC(O)CH_3), 48.7 (d, C-2), 61.4 (t, C-6), 66.7 (d, C-4), 71.6 (d, C-3), 75.3 (d, C-5), 82.1 (d, C-1), 170.1 (s, C(O)CH_3), 170.2 (s, C(O)CH_3), 170.4 (s, C(O)CH_3), 170.8 (s, C(O)CH_3), 193.7 (s, SC(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{16}H_{24}NO_9S^+$: 406.1172. Found: 406.1165 (MH^+).

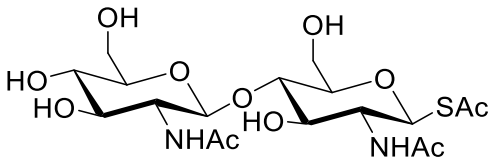
1-Thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-α-D-mannopyranoside 3.14



N-Acetyl-D-mannosamine (500 mg, 2.26 mmol) and triethylamine (3.1 mL, 22.6 mmol) were stirred in D_2O (10 mL) and the mixture was cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (1.2 g, 6.78 mmol) was added. After 0.5 h, thioacetic acid (2.4 mL, 33.9 mmol) was added dropwise, and the reaction mixture was allowed to stir at 0 °C for an additional 0.5 h. The reaction mixture was then diluted with water (10 mL) and washed with CH_2Cl_2 (5 x 20 mL). The aqueous layer was concentrated *in vacuo* and the residue stirred in pyridine (7 mL) under an atmosphere of nitrogen. Acetic anhydride (7 mL, 74.1 mmol) was added and the mixture was stirred for 12 h, after which time t.l.c. (petrol:ethyl acetate, 2:1) indicated the formation of a single major product (R_f 0.8). Water (10 mL) was added, and the mixture was then washed with CH_2Cl_2 (3 x 100 mL). The combined organic extracts were washed with aqueous aqueous HCl (1 M, 100 mL), $NaHCO_3$ (sat. aqueous soln., 100 mL), water (100 mL) and brine (100 mL), dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The residue was purified by flash column chromatography (ethyl acetate:petrol, 3:2) to afford 1-thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-α-D-mannopyranoside **3.14** (0.807 g,

88%) as a colourless oil; $[\alpha]_D^{20}$ +55 (*c*, 1.0 in $CHCl_3$); υ_{max} (neat) 1740 cm^{-1} (OC=O), 1711 cm^{-1} (SC=O), 1684 cm^{-1} (NHC=O); δ_H (400 MHz, $CDCl_3$) 1.99 (3H, s, C(O)C$\underline{H}_3$), 2.05 (3H, s, C(O)C$\underline{H}_3$), 2.06 (3H, s, C(O)C$\underline{H}_3$), 2.08 (3H, s, C(O)C$\underline{H}_3$), 2.41 (3H, s, SC(O)C$\underline{H}_3$), 3.96 - 4.06 (2H, m, H-5 & H-6), 4.28 (1H, dd, $J_{5,6'}$ 4.7 Hz, $J_{6,6'}$ 12.1, H-6'), 4.65 - 4.70 (1H, m, H-2), 5.04 (1H, dd, $J_{3,4}$ 9.8 Hz, $J_{2,3}$ 3.9 Hz, H-3), 5.15 (1H, t, *J* 9.8 Hz, H-4), 5.85 (1H, d, $J_{1,2}$ 1.6 Hz, H-1), 6.02 (1H, d, *J* 8.6 Hz, N$\underline{H}$C(O)CH_3); δ_C (100.5 MHz, $CDCl_3$) 20.6 (q, C(O)$\underline{C}H_3$), 20.6 (q, C(O)$\underline{C}H_3$), 20.7 (q, C(O)$\underline{C}H_3$), 23.2 (q, C(O)$\underline{C}H_3$), 31.2 (q, SC(O)$\underline{C}H_3$), 51.4 (d, C-2), 62.1 (t, C-6), 65.7 (d, C-4), 70.1 (d, C-3), 72.0 (d, C-5), 81.1 (d, C-1), 169.7 (s, $\underline{C}$(O)CH_3), 169.9 (s, $\underline{C}$(O)CH_3), 170.0 (s, $\underline{C}$(O)CH_3), 170.5 (s, $\underline{C}$(O)CH_3), 190.8 (s, S$\underline{C}$(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{16}H_{24}NO_9S^+$: 406.1172. Found: 406.1152 (MH^+).

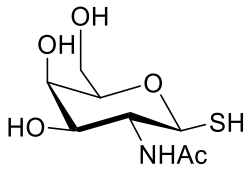
1-Thioacetyl-*N,N'*-diacetylchitobiose 3.15



N,N'-Diacetylchitobiose (9.8 mg, 0.023 mmol) and triethylamine (32 µL, 0.23 mmol) were stirred in D_2O (92 µL) and cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (11.8 mg, 0.069 mmol) was added. After 0.5 h, thioacetic acid (24 µL, 0.345 mmol) was added dropwise and the reaction mixture was allowed to stir for an additional 1 h at 0 °C. The reaction was purified by HPLC (column: Luna 5u C18 (100 Å) column (Phenomenex); gradient: 100% water for 15 min, followed by an increase to 100% MeOH over 5 min; column oven: 40 °C; flow rate: 2 mL/min; detection: UV 210 nm) to afford 1-thioacetyl-*N,N'*-diacetylchitobiose **3.15** (10.3 mg, 93%) as a white powder; υ_{max} (neat) 3270 cm^{-1} (OH), 1697 cm^{-1} (SC=O), 1649 cm^{-1} (NHC=O); $[\alpha]_D$ +0.1 (*c*, 0.5 in water); δ_H (400 MHz, D_2O) 1.97 (3H, s, NHC(O)C$\underline{H}_3$), 2.05 (3H, s, NHC(O)C$\underline{H}_3$), 2.39 (3H, s, SC(O)C$\underline{H}_3$), 3.43 - 3.50 (2H, m, H-3_a &

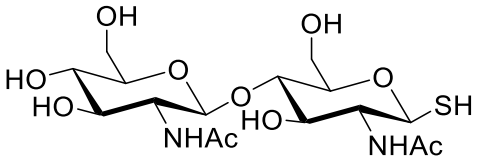
H-3$_b$), 3.55 (1H, t, *J* 9.1 Hz, H-4$_b$), 3.58 - 3.68 (3H, m, H-4$_a$, H-5$_a$ & H-6$_a$), 3.69 - 3.84 (4H, m, H-2$_b$, H-5$_b$, H-6'$_a$ & H-6$_b$), 3.86 - 3.96 (2H, H-2$_a$ & H-6'$_b$), 4.57 (1H, d, *J* 8.6 Hz, H-1$_b$), 5.18 (1H, d, *J* 11.0 Hz, H-1$_a$); δ_C (100.5 MHz, D_2O) 21.9 (q, NHC(O)$\underline{C}H_3$), 22.0 (q, NHC(O)$\underline{C}H_3$), 30.1 (q, SC(O)$\underline{C}H_3$), 52.8 (d, C-2$_a$), 55.5 (d, C-2$_b$), 59.9 (t, C-6$_a$), 60.5 (t, C-6$_b$), 69.6 (d, C-3$_a$), 73.4 (d, C-4$_b$), 73.5 (d, C-5$_b$), 75.8 (d, C-3$_b$), 78.6, 78.9 (2 x d, C-4$_a$ & C-5$_a$), 80.8 (d, C-1$_a$), 101.3 (d, C-1$_b$), 174.5 (s, 2 x NH$\underline{C}$(O)$\underline{C}H_3$), 197.4 (s, S$\underline{C}$(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{18}H_{31}N_2O_{11}S^+$: 483.1643. Found: 483.1655 (MH^+).

1-Thio-2-acetamido-2-deoxy-β-D-galactopyranose 3.16



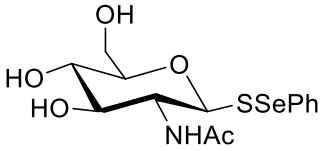
Sodium metal (25.4 mg, 1.1 mmol) was added to MeOH (5 mL) and stirred. 1-Thioacetyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-β-D-galactopyranoside **3.13** (179 mg, 0.44 mmol) was added and the mixture was stirred for 0.5 h, at which point t.l.c. (ethyl acetate:petrol, 2:1) indicated complete consumption of starting material (R_f 0.1) and formation of a single product (R_f 0). Dowex® 50WX8 (H^+) ion exchange resin was added portion-wise until the reaction reached neutral pH. The reaction was then filtered and concentrated *in vacuo* to afford 1-thio-2-acetamido-2-deoxy-β-D-galactopyranose **3.16** (108 mg, quant.) as a white amorphous solid; $[\alpha]_D^{20}$ +29 (*c*, 0.9 in MeOH); υ_{max} (neat) 3275 cm^{-1} (b, OH, NH), 2545 cm^{-1} (SH), 1625 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.94 (3H, s, C(O)C$\underline{H}_3$), 3.56 - 3.67 (4H, m, H-3, H-5, H-6 & H-6'), 3.80 - 3.88 (2H, m, H-2 & H-4), 4.50 (1H, d, $J_{1,2}$ 10.2 Hz, H-1); δ_C (100.5 MHz, D_2O) 22.2 (q, C(O)$\underline{C}H_3$), 54.6 (d, C-2), 61.1 (t, C-6), 67.8 (d, C-4), 71.8 (d, C-3), 79.3 (d, C-1), 79.3 (d, C-5), 174.8 (s, $\underline{C}$(O)CH_3); HRMS (ESI-TOF): calcd. for $C_8H_{15}NO_5S^+$: 238.0744. Found: 238.0745 (MH^+).

1-Thio-*N,N'*-diacetylchitobiose 3.18



1-Thioacetyl-*N,N'*-diacetylchitobiose **3.15** (2.2 mg, 4.6 µmol) was dissolved in aqueous NaOH (1 M, 1.5 mL) and stirred for 1 h. Amberlite® IR-120 (H^+) was then added until pH 7 was reached. The solution was then filtered and lyophilised to afford 1-thio-*N,N'*-diacetylchitobiose **3.18** (2.0 mg, quant.) as a white powder; $[\alpha]_D^{20}$ -5 (*c*, 0.28 in water); δ_H (400 MHz, D_2O)[12] 2.02 (3H, s, NHC(O)C$\underline{H}_3$), 2.04 (3H, s, NHC(O)C$\underline{H}_3$), 3.42 - 3.57 (4H, m, H-3_a, H-4_a, H-5_a & H-3_b), 3.57 - 3.67 (3H, m, H-4_b, H-5_b & H-6_a), 3.68 - 3.76 (3H, m, H-2_a, H-2_b & H-6_b), 3.81 (1H, dd, $J_{5,6}$ 1.2 Hz, $J_{6,6'}$ 12.1 Hz, H-6_a'), 3.89 (1H, dd, $J_{5,6'}$ 12.3 Hz, $J_{6,6'}$ 1.6 Hz, H-6_b'), 4.56 (1H, d, $J_{1b,2b}$ 8.2 Hz, H-1_b), 4.65 (1H, d, $J_{1a,2a}$ 9.8 Hz, H-1_a); δ_C (100.5 MHz, D_2O) 22.0 (q, NHC(O)$\underline{C}H_3$), 22.1 (q, NHC(O)$\underline{C}H_3$), 55.5 (d, C-2_b), 57.2 (d, C-2_a), 60.1 (t, C-6), 60.5 (t, C-6), 69.6 (d, C-3_a), 73.4 (d, C-3_b), 73.5 (d, C-5_a), 75.8 (d, C-5_b), 78.6 (d, C-4_a), 78.8 (d, H-1_a), 79.0 (d, C-4_b), 101.3 (d, H-1_b), 174.5 (s, 2 x q, NH$\underline{C}$(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{16}H_{29}N_2O_{10}S^+$: 441.1537. Found: 441.1544 (MH^+).

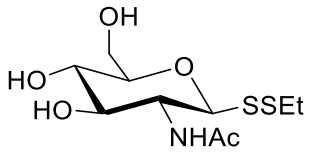
Phenyl 2-acetamido-2-deoxy-1-selenylsulfide-β-D-glucopyranoside 3.20[7]



Phenylselenyl bromide (216 mg, 0.92 mmol) was dissolved in 1,4-dioxane (4 mL) and stirred at room temperature under and atmosphere of nitrogen. 1-Thio-2-acetamido-2-deoxy-β-D-glucopyranose **3.5** (200 mg, 0.84 mmol) was dissolved in MeOH (15 mL) and added dropwise to the reaction mixture, which was then stirred for 2 min. At this point, t.l.c. (ethyl acetate:MeOH 9:1) indicated the formation of a major product (R_f 0.4). The reaction was

concentrated *in vacuo*, pre-absorbed onto silica and purified by flash column chromatography (ethyl acetate:MeOH, 9:1) to afford phenyl 2-acetamido-2-deoxy-1-selenylsulfide-β-D-glucopyranoside **3.20** (210 mg, 64%) as a white amorphous solid; $[\alpha]_D^{20}$ -109 (*c*, 1.1 in MeOH) [lit. $[\alpha]_D^{22}$ -174 (*c*, 1.0 in MeOH)][7]; υ_{max} (neat) 3258 cm^{-1} (OH, NH), 1652 cm^{-1} (NHC=O); δ_H (400 MHz, CD_3OD)[7] 1.94 (3H, s, C(O)C<u>H</u>$_3$), 3.28 - 3.39 (2H, m, H-4 & H-5), 3.44 - 3.54 (1H, m, H-3), 3.63 (1H, dd, $J_{6,6'}$ 12.1 Hz, $J_{5,6}$ 5.1 Hz, H-6), 3.78 - 3.90 (2H, m, H-2 & H-6'), 4.63 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 7.23 - 7.33 (3H, m, 3 x Ar-H), 7.68 - 7.73 (2H, m, 2 x Ar-H); δ_C (100.5 MHz, CD_3OD) 21.5 (q, C(O)<u>C</u>H$_3$), 55.8 (d, C-2), 61.6 (t, C-6), 70.4 (d, C-4), 75.6 (d, C-3), 80.9 (d, C-5), 89.1 (d, C-1), 127.2, 128.7, 130.4 (3 x d, 3 x Ar-C), 132.6 (s, Ar-C), 172.1 (s, <u>C</u>(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{14}H_{20}NO_5S^{80}Se^+$: 394.0222. Found: 394.0217 (MH^+).

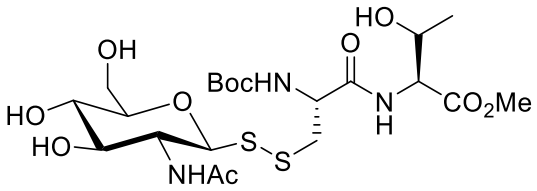
Ethyl 2-acetamido-2-deoxy-1-disulfide-β-D-glucopyranoside 3.21[7]



Phenyl 2-acetamido-2-deoxy-1-selenylsulfide-β-D-glucopyranoside **3.20** (140 mg, 0.4 mmol) was dissolved in MeOH (10 mL) and stirred under an atmosphere of nitrogen. A solution of ethanethiol (10 µL, 0.13 mmol) and triethylamine (55 µL, 0.4 mmol) in MeOH (5 mL) was added dropwise to the reaction over 1 h. After 1 h, t.l.c. (ethyl acetate:MeOH, 9:1) indicated the formation of a major product (R_f 0.2). The mixture was concentrated *in vacuo* and the residue was purified by flash column chromatography (ethyl acetate:MeOH, 9:1) to afford ethyl 2-acetamido-2-deoxy-1-disulfide-β-D-glucopyranoside **3.21** (36.7 mg, 95%) as a white amorphous solid; υ_{max} (neat)[7] 3271 cm^{-1} (b, OH, NH), 1650 cm^{-1} (C=O); δ_H (400 MHz, CD_3OD)[7] 1.29 (3H, t, *J* 7.4 Hz, SCH$_2$C<u>H</u>$_3$), 1.96 (3H, s, C(O)C<u>H</u>$_3$), 2.76 - 2.88 (2H, m,

$SC\underline{H}_2CH_3$), 3.29 - 3.38 (2H, m, H-4 & H-5), 3.44 - 3.51 (1H, m, H-3), 3.64 - 3.70 (1H, dd, $J_{5,6}$ 5.5 Hz, $J_{6,6'}$ 12.1 Hz , H-6), 3.79 - 3.90 (2H, m, H-2 & H-6'), 4.53 (1H, d, $J_{1,2}$ 10.6 Hz, H-1); δ_C (100.5 MHz, CD_3OD) 13.4 (q, $SCH_2\underline{C}H_3$), 21.5 (q, $C(O)\underline{C}H_3$), 33.2 (t, $S\underline{C}H_2CH_3$), 54.5 (d, C-2), 61.6 (t, C-6), 70.4 (d, C-4), 75.8 (d, C-3), 81.1 (d, C-5), 91.1 (d, C-1), 172.0 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{10}H_{19}NO_5S_2^+$: 298.0777. Found: 298.0778 (MH^+).

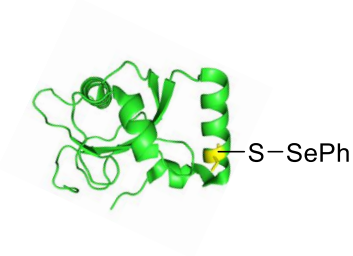
_N_-Butoxycarbonyl-L-cysteinyl-(_S_-2-acetamido-2-deoxy-1-β-D-glucopyranosyl disulfide)-L-threonine methyl ester 3.23[7]



Phenyl 2-acetamido-2-deoxy-1-selenylsulfide-β-D-glucopyranoside **3.20** (70 mg, 0.2 mmol) was dissolved in MeOH (8 mL) and stirred under an atmosphere of nitrogen. A solution of *N*-butoxycarbonyl-L-cysteine-L-threonine methyl ester **3.22** (22 mg, 0.07 mmol) and triethylamine (10 µL, 0.1 mmol) in MeOH (5 mL) was added dropwise. After 10 min, t.l.c. (ethyl acetate:MeOH, 9:1) indicated the formation of a major product (R_f 0.4). The reaction mixture was concentrated *in vacuo* and the residue was purified by flash column chromatography (ethyl acetate:MeOH, 9:1) to afford *N*-butoxy-L-cysteinyl-(*S*-2-acetamido-2-deoxy-1-β-D-glucopyranosyl disulfide)-L-threonine methyl ester **3.23** (30.8 mg, 77%) as a white amorphous solid; υ_{max} (neat)[7] 3331 cm^{-1} (b, OH, NH), 1735 cm^{-1} (C=O), 1655 cm^{-1} (C(O)NH); δ_H (400 MHz, CD_3OD)[7] 1.17 (3H, d, *J* 6.3 Hz, $CHC\underline{H}_3$), 1.46 (9H, s, $C(C\underline{H}_3)_3$), 1.96 (3H, s, $C(O)C\underline{H}_3$), 2.94 (1H, dd, $J_{CH,H}$ 13.5 Hz, $J_{CH,\alpha H}$ 9.2 Hz, $C\underline{H}HCys$), 3.33 - 3.44 (3H, m, $CH\underline{H}Cys$, H-4 & H-5), 3.51 (1H, t, *J* 9.0 Hz, H-3), 3.75 (3H, s, OMe), 3.87 - 4.01 (2H, m, H-2 & H-6), 4.26 - 4.35 (1H, m, $C\underline{H}CH_3$), 4.46 (1H, d, *J* 3.1 Hz, αHThr), 4.57 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 4.59 - 4.68 (1H, m, αHCys); δ_C (100.5 MHz, CD_3OD) 18.9 (q, $CH\underline{C}H_3$),

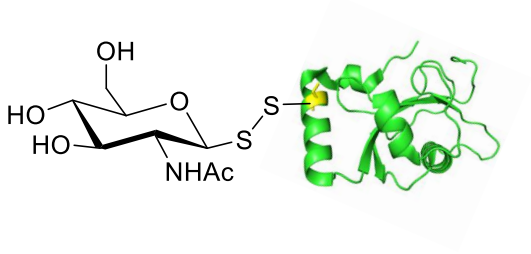
21.7 (q, C(O)<u>C</u>H$_3$), 27.3 (C(<u>C</u>H$_3$)$_3$), 41.1 (t, <u>C</u>H$_2$), 51.6 (q, O<u>C</u>H$_3$), 53.9 (d, αCCys), 54.0 (d, C-2), 57.9 (d, αCThr), 61.2 (t, C-6), 67.1 (d, <u>C</u>HCH$_3$), 70.1 (d, C-4), 75.7 (d, C-3), 79.7 (s, <u>C</u>H(CH$_3$)$_3$), 81.0 (d, C-5), 89.3 (d, H-1), 156.4, 172.1, 172.2, 172.7 (4 x d, 4 x C(O)); HRMS (ESI-TOF): calcd. for $C_{21}H_{38}N_3O_{11}S_2^+$: 572.1942. Found: 572.1957 (MH^+).

GABARAPL2-Cys-17-SePh 3.24



A solution of GABARAPL2 (0.3 mg) in CHES buffer (0.5 mL, 70 mM, pH 9.5) was prepared. Phenylselenyl bromide (2.6 mg, 0.011 mmol) was dissolved in MeCN (200 µL), of which 20 µL was added to the GABARAPL2 solution. The mixture was stirred for 1 h, then filtered through a 0.2 µm filter and dialysed against water (2 x 500 mL) to afford GABARAPL2-Cys-17-SePh **3.24**; HRMS (ESI-TOF): calcd. for (GABARAPL2-Cys-17-SePh)K^+: 14005. Found: 14005 (MK^+).

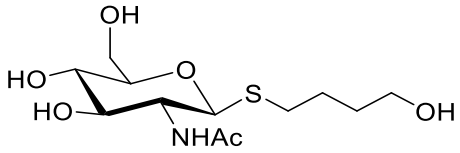
GABARAPL2-Cys-17-SGlcNAc 3.25



A solution of 1-thio-2-acetamido-2-deoxy-β-D-glucopyranose **3.5** (1 mg) in water (200 µL) was prepared and 50 µL was added to a solution of GABARAPL2-Cys-17-SePh **3.24** (0.3 mg) in CHES buffer (1.0 mL, 70 mM, pH 9.5). The reaction was stirred for 1 h and then purified

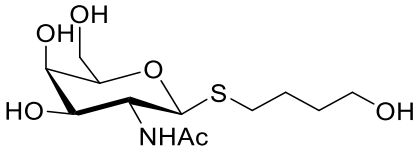
using a Sephadex G-25 HiTrap^{TM} desalting column to afford GABARAPL2-Cys-17-SGlcNAc **3.25**; HRMS (ESI-TOF): calcd. for (GABARAPL2-Cys-17-SGlcNAc)K^+: 14084. Found: 14084 (MK^+).

4-Hydroxybutyl 1-thio-2-deoxy-2-acetamido-β-D-glucopyranoside 3.26



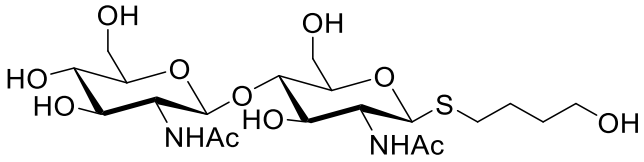
1-Thio-2-deoxy-2-acetamido-β-D-glucopyranoside **3.5** (200 mg, 0.84 mmol) and 3-buten-1-ol (65 µL, 0.76 mmol) were dissolved in water (7 mL). 2,2-Dimethoxy-2-phenylacetophenone (19.6 mg, 0.076 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 2 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by flash column chromatography ($CHCl_3$:MeOH 5:1) to afford 4-hydroxybutyl 1-thio-2-deoxy-2-acetamido-β-D-glucopyranoside **3.26** (142.4 mg, 61%) as a colourless solid; $[\alpha]_D^{20}$ -29 (*c*, 1.0 in MeOH); υ_{max} (neat) 1620 cm^{-1} (C=O); δ_H (400 MHz, D_2O) 1.43 - 1.58 (4H, m, 2 x C$\underline{H}_2$), 1.90 (3H, s, C(O)C$\underline{H}_3$) 2.51 - 2.71 (2H, m, C$\underline{H}_2$S), 3.29-3.34 (2H, m, H-4 & H-5), 3.36 - 3.44 (1H, m, H-3), 3.47 (2H, t, *J* 6.1 Hz, C$\underline{H}_2$OH), 3.55 - 3.65 (2H, m, H-2 & H-6), 3.77 (1H, d, $J_{6,6'}$ 12.5 Hz, H-6'), 4.47 (1H, d, $J_{1,2}$ 10.2 Hz, H-1); δ_C (100.5 MHz, D_2O) 22.1 (t, C(O)$\underline{C}H_3$), 25.5 (t, $\underline{C}H_2CH_2OH$), 29.9 (t, $\underline{C}H_2S$), 30.3 (t, $SCH_2\underline{C}H_2$), 54.7 (d, C-2), 60.8 (t, C-6), 61.1 (t, $\underline{C}H_2OH$), 69.7 (d, C-4), 75.1 (d, C-3), 79.8 (d, C-5), 84.2 (d, C-1), 174.3 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{12}H_{23}NO_6SNa^+$: 332.1138. Found: 332.1146 (MNa^+).

4-Hydroxybutyl 1-thio-2-deoxy-2-acetamido-β-D-galactopyranoside 3.27



1-Thio-2-deoxy-2-acetamido-β-D-galactopyranoside **3.16** (131.3 mg, 0.55 mmol) and 3-buten-1-ol (43 µL, 0.50 mmol) were dissolved in water (5 mL). 2,2-Dimethoxy-2-phenylacetophenone (12.9 mg, 0.050 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 2 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by flash column chromatography ($CHCl_3$:MeOH 5:1) to afford 4-hydroxybutyl 1-thio-2-deoxy-2-acetamido-β-D-galactopyranoside **3.27** (95 mg, 61%) as a colourless solid; $[\alpha]_D^{20}$ +2 (*c*, 1.0 in MeOH); υ_{max} (neat) 1627 cm^{-1} (C=O); δ_H (400 MHz, D_2O) 1.56 - 1.69 (4H, m, 2 x C$\underline{H}_2$), 2.01 (3H, s, C(O)C$\underline{H}_3$), 2.64 - 2.81 (2H, m, C$\underline{H}_2$S), 3.59 (2H, t, *J* 6.1 Hz, C$\underline{H}_2$OH), 3.64 - 3.79 (4H, m, H-3, H-5, H-6 & H-6'), 3.89 - 3.95 (1H, m, H-2), 3.96 (1H, d, $J_{4,5}$ 3.1 Hz, H-4), 4.52 (1H, d, $J_{1,2}$ 10.6 Hz, H-1); δ_C (100.5 MHz, D_2O) 22.1 (q, C(O)$\underline{C}H_3$), 25.5 (t, $\underline{C}H_2CH_2S$), 29.9 ($\underline{C}H_2S$), 30.3 ($\underline{C}H_2CH_2OH$), 51.3 (d, C-2), 61.0 (t, C-6), 61.1 (t, $\underline{C}H_2OH$), 67.8 (d, C-4), 72.0 (d, C-3), 78.9 (d, C-5), 84.6 (d, C-1), 174.5 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{12}H_{23}NO_6SNa^+$: 332.1138. Found: 332.1141 (MNa^+).

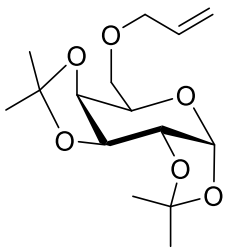
4-Hydroxybutyl 1-thio-*N,N'*-diacetylchitobiose 3.28



1-Thio-*N,N'*-diacetylchitobiose **3.18** (7.7 mg, 0.017 mmol) and 3-buten-1-ol (1.5 µL, 0.017 mmol) were dissolved in water (200 µL). 2,2-Dimethoxy-2-phenylacetophenone (0.5 mg, 0.002 mmol) was added, and the reaction mixture was stirred and irradiated in a

Rayonet photoreactor (wavelength: 254 nm) for 2 h. The reaction was purified by HPLC (column: Jupiter 10u C18 (90 Å) column (Phenomenex); gradient: 100% water for 15 min, followed by an increase to 100% MeOH over 5 min; column oven: 40 °C; flow rate: 2 mL/min; detection: UV 210 nm) to afford 4-hydroxybutyl 1-thio-*N,N'*-diacetylchitobiose **3.28** (5.6 mg, 64%) as a white powder; υ_{max}(neat) 3288 cm^{-1} (OH), 1643 cm^{-1} (NHC=O); $[\alpha]_D^{20}$ -13 (*c*, 0.3 in H_2O); δ_H (400 MHz, D_2O) 1.45-1.58 (4H, m, 2 x C<u>H</u>$_2$), 1.91 (3H, s, C(O)C<u>H</u>$_3$), 1.94 (3H, s, C(O)C<u>H</u>$_3$), 2.50 - 2.70 (2H, m, C<u>H</u>$_2$S), 3.33 - 3.67 (12H, m), 3.72 (1H, dd, $J_{5,6'}$ 0.8 Hz, $J_{6,6'}$ 12.1 Hz, H-6_a'), 3.80 (1H, d, $J_{6,6'}$ 11.7 Hz, H-6_b'), 4.42 - 4.53 (2H, m, H-1_a & H-1_b); δ_C (100.5 MHz, D_2O) 22.0 (q, C(O)<u>C</u>H$_3$), 22.1 (q, C(O)<u>C</u>H$_3$), 25.5 (t, SCH$_2$<u>C</u>H$_2$), 30.0 (t, S<u>C</u>H$_2$), 30.3 (t, <u>C</u>H$_2$CH$_2$OH), 54.2 (d, C-2_a), 55.5 (d, C-2_b), 60.2 (t, C-6_a), 60.5 (t, C-6_b), 61.1 (t, O<u>C</u>H$_2$), 69.6, 73.4, 73.7, 75.8, 78.4, 79.2 (6 x d), 84.2 (C-1_a), 101.4 (C-1_b), 174.3 (s, <u>C</u>(O)CH$_3$), 174.5 (s, <u>C</u>(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{20}H_{37}N_2O_{11}S^+$: 513.2113. Found: 513.2124 (MH^+).

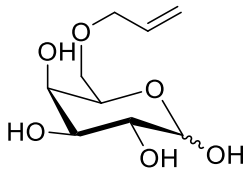
6-*O*-Allyl-1,2:3,4-di-*O*-isopropylidene-α-D-galactopyranoside 3.30[13]



1,2:3,4-Di-*O*-isopropylidene-α-D-galactopyranoside **3.29** (1.185 g, 4.55 mmol) and allyl bromide (0.86 mL, 10 mmol) were dissolved in dry MeCN (5 mL) under an atmosphere of nitrogen. Potassium hydroxide (0.56 g, 10 mmol) was suspended in the reaction mixture and the reaction was stirred for 21 h, at which point t.l.c. (petrol:ethyl acetate 3:1) indicated complete consumption of starting material (R_f 0.1) and formation of a single product (R_f 0.6). The reaction was concentrated *in vacuo*, the residue dissolved in CH_2Cl_2 (15 mL), and washed

with water (15 mL). The aqueous phase was washed with CH_2Cl_2 (3 x 15 mL) and the combined organic extracts were dried ($MgSO_4$) and concentrated *in vacuo*. The product was purified by flash column chromatography (petrol:ethyl acetate 3:1) to afford 6-*O*-allyl-1,2:3,4-di-*O*-isopropylidene-α-D-galactopyranoside **3.30** (1.28 g, 94%) as a colourless oil; $[\alpha]_D^{20}$ -69 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D$ -71.8 (*c*, 1.0 in $CHCl_3$)][14]; δ_H (400 MHz, $CDCl_3$)[15] 1.31 (3H, s, C$\underline{H}_3$), 1.32 (3H, s, C$\underline{H}_3$), 1.43 (3H, s, C$\underline{H}_3$), 1.52 (3H, s, C$\underline{H}_3$), 3.53 - 3.67 (2H, m, C$\underline{H}_2$CH=CH$_2$), 3.93 - 3.99 (1H, m, H-5), 4.00 - 4.05 (2H, m, H-6 & H-6'), 4.25 (1H, dd, $J_{3,4}$ 8.0, $J_{4,5}$ 1.8 Hz, H-4), 4.29 (1H, dd, $J_{1,2}$ 4.9 Hz, $J_{2,3}$ 2.5 Hz, H-2), 4.58 (1H, dd, $J_{2,3}$ 2.3 Hz, $J_{3,4}$ 7.8 Hz, H-3), 5.13 - 5.18 (1H, m, CH$_2$CH=C$\underline{H}_a$H$_b$), 5.26 (1H, dq, *J* 17.2 Hz, *J* 1.6 Hz, CH$_2$CH=CH$_a\underline{H}_b$), 5.52 (1H, d, $J_{1,2}$ 4.7 Hz, H-1), 5.90 (1H, ddt, *J* 17.2 Hz, *J* 10.4 Hz, *J* 5.8 Hz, CH$_2$C$\underline{H}$=CH$_2$); δ_C (100.5 MHz, $CDCl_3$) 24.4 (q, C$\underline{H}_3$), 24.9 (q, C$\underline{H}_3$), 25.9 (q, C$\underline{H}_3$), 26.0 (q, C$\underline{H}_3$), 66.8 (d, C-5), 68.8 (t, C$\underline{H}_2$CH=CH$_2$), 70.5 (d, C-2), 70.6 (d, C-3), 71.2 (d, C-4), 72.3 (t, C-6), 96.3 (d, C-1), 108.5 (s, C$\underline{(CH_3)_2}$), 109.2 (s, C$\underline{(CH_3)_2}$), 117.0 (t, CH$_2$CH=C$\underline{H}_2$), 134.8 (d, CH$_2$C$\underline{H}$=CH$_2$); HRMS (ESI-TOF): calcd. for $C_{15}H_{24}O_6Na^+$: 323.1465. Found: 323.1469 (MNa^+).

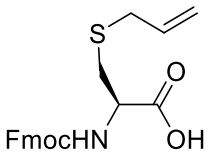
6-*O*-Allyl-D-galactopyranoside 3.31[13]



6-*O*-Allyl-1,2:3,4-di-*O*-isopropylidene-α-D-galactopyranoside **3.30** (1.12 g, 3.73 mmol) was dissolved in water (50 mL). Amberlite® IR-120 (H^+, 4.0 g) was added, and the reaction was heated to 80 °C and stirred for 16 h. The reaction was then cooled to RT and lyophilised to afford 6-*O*-allyl-D-galactopyranoside **3.31** (0.84 g, quant.) as a colourless oil (1:1.3 α:β); δ_H (400 MHz, D_2O) 3.46 (1H, dd, $J_{1,2}$ 7.8 Hz, $J_{3,4}$ 9.8 Hz, H-2β), 3.62 (1H, dd, $J_{2,3}$ 10.2 Hz, $J_{3,4}$ 3.5 Hz, H-3β), 3.64 - 3.67 (3.5H, m, H-6α, H-6'α, H-6β & H-6'β), 3.75 - 3.85 (2.5H, m,

H-2α, H-3α & H-5β), 3.88 (1H, d, $J_{3,4}$ 3.5 Hz, H-4β), 3.94 (0.8H, d, $J_{3,4}$ 3.1 Hz, H-4α), 4.03 - 4.09 (3.6H, m, 2 x C$\underline{H}_2$CH=CH$_2$), 4.19 (0.8H, t, *J* 6.3 Hz, H-5α), 4.55 (1H, d, $J_{1,2}$ 7.8 Hz, H-1β), 5.23 (0.8H, d, $J_{1,2}$ 3.9 Hz, H-1α), 5.26 (1.8H, d, *J* 10.6 Hz, 2 x CH$_2$CH=C$\underline{H}_a$H$_b$), 5.32 (1.8H, d, *J* 17.2 Hz, 2 x CH$_2$CH=CH$_a\underline{H}_b$), 5.86 - 6.01 (1.8H, m, 2 x CH$_2$C$\underline{H}$=CH$_2$); δ_C (100.5 MHz, D$_2$O) 68.2 (d, C-2α), 68.6 (d, C-3α), 68.9 (d, C-5α), 69.0 (d, C-4β), 69.2, 69.4 (2 x t, C-6α & C-6β), 69.5 (d, C-4α), 71.7 (d, C-2β), 71.9, 72.0 (2 x t, $\underline{C}$H$_2$CH=CH$_2$α & $\underline{C}$H$_2$CH=CH$_2$β), 72.6 (d, C-3β), 73.3 (d, C-5β), 92.2 (d, C-1α), 96.3 (d, C-1β), 118.5 (t, CH$_2$CH=$\underline{C}$H$_2$α & CH$_2$CH=$\underline{C}$H$_2$β), 133.5, 133.5 (2 x d, CH$_2\underline{C}$H=CH$_2$α & CH$_2\underline{C}$H=CH$_2$β); HRMS (ESI-TOF): calcd. for $C_9H_{16}O_6Na^+$: 243.0839. Found: 243.0840 (MNa^+).

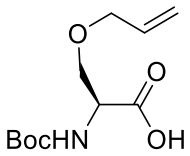
_N_-[(9-Fluorenylmethoxy)carbonyl]-_S_-allyl-L-cysteine 3.33[16]



An aqueous solution of ammonium hydroxide (2 M, 18 mL, 21 mmol) was added to EtOH (15 mL), followed by L-cysteine monohydrate hydrochloride **3.32** (1.05 g, 6 mmol). Allyl bromide (0.78 mL, 9 mmol) was then added and the reaction was stirred for 20 h, at which point t.l.c. ($CHCl_3$:MeOH 5:1) indicated complete consumption of starting material (R_f 0) and formation of a single product (R_f 0.1). The reaction was concentrated *in vacuo* and the white solid produced was washed with EtOH (2 x 10 mL). The crude solid was dissolved in water (15 mL), cooled to 0 °C, and $NaHCO_3$ (1.01 g, 12 mmol) was added. FmocCl (2.33 g, 9 mmol) was dissolved in 1,4-dioxane (20 mL) and the resulting solution was then added to the reaction mixture. The reaction mixture was allowed to warm to RT, and stirred for 20 h. The reaction was then diluted with water (15 mL), acidified to pH 1 by the dropwise addition of H_2SO_4 (conc.), and extracted with ethyl acetate (4 x 30 mL). The combined organic extracts were concentrated *in vacuo* and the residue produced was purified by flash column chromatography

(CH_2Cl_2:^{i}PrOH:AcOH 97:2:1) to afford *N*-[(9-fluorenylmethoxy)carbonyl]-*S*-allyl-L-cysteine **3.33** (103 mg, 5%) as a white solid; $[\alpha]_D^{20}$ +12 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D$ -12.2 (*c*, 1.0 in $CHCl_2$)][16]; υ_{max} (neat) 1692 cm^{-1} (C=O); δ_H (400 MHz, $CDCl_3$)[16] 2.88 - 3.04 (2H, m, Cys-H_β), 3.15 (2H, d, *J* 6.7 Hz, C$\underline{H}_2$CH=CH_2), 4.23 (1H, t, *J* 6.8 Hz, Fmoc C$\underline{H}$$CH_2$), 4.42 (2H, d, *J* 6.3 Hz, Fmoc CHC$\underline{H}_2$), 4.56 - 4.63 (1H, m, Cys-H_α), 5.11 (2H, d, *J* 10.6 Hz, CH_2CH=C$\underline{H}_2$), 5.57 (1H, d, *J* 7.0 Hz, N$\underline{H}$), 5.67 - 5.82 (1H, m, CH_2C$\underline{H}$=CH_2), 7.31 (2H, t, *J* 7.4 Hz, 2 x Ar-H), 7.39 (2H, t, *J* 7.4 Hz, 2 x Ar-H), 7.59 (2H, d, *J* 6.3 Hz, 2 x Ar-H), 7.76 (2H, d, *J* 7.8 Hz, 2 x Ar-H); δ_C (100.5 MHz, $CDCl_3$)[16] 32.5 (t, Cys-C_β), 35.3 (t, $\underline{C}H_2$CH=CH_2), 47.1 (Fmoc $\underline{C}$HCH_2), 52.4 (d, Cys-C_β), 67.3 (t, Fmoc CH$\underline{C}H_2$), 118.1 (t, CH_2CH=$\underline{C}H_2$), 120.0, 125.0, 127.1, 127.7 (4 x d, 4 x Ar-C), 133.5, 141.3 (2 x s, 2 x Ar-C), 143.6; HRMS (ESI-TOF): calcd. for $C_{21}H_{22}NO_4S^+$: 384.1264. Found: 384.1271 (MH^+).

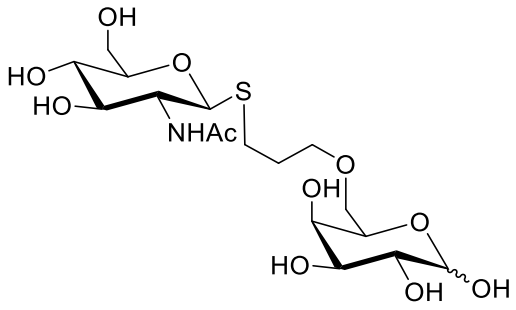
_N_-_tert_-Butoxycarbonyl-_O_-allyl-L-serine 3.35[17]



L-Serine **3.34** (500 mg, 4.76 mmol) was dissolved in a mixture of 1,4-dioxane (12 mL) and aqueous NaOH (1 M, 5.7 mL) at 0 °C. Di-*tert*-butyl dicarbonate (1.25 g, 5.71 mmol) was added portionwise to the reaction, and the reaction was then warmed to RT and stirred for 24 h. The reaction was concentrated *in vacuo* to remove the 1,4-dioxane and the aqueous layer was washed with Et_2O (10 mL), acidified to pH 2-3 with H_2SO_4 (1 M, 10 mL), and extracted with ethyl acetate (3 x 10 mL). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was dissolved in anhydrous DMF (30 mL) and NaH (60% dispersion in mineral oil, 460 mg, 11.4 mmol) was added at 0 °C under an atmosphere of nitrogen. Allyl bromide (0.49 mL, 5.71 mmol) was added dropwise to the mixture and the reaction was stirred at RT for 5 h. The reaction was concentrated *in vacuo* and the residue was

partitioned between ethyl acetate (50 mL) and water (50 mL). The aqueous layer was acidified with HCl (1 M, 20 mL) and extracted with ethyl acetate (3 x 50 mL). The combined organic extracts were washed with water (50 mL), followed by brine (50 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated *in vacuo* to afford *N*-*tert*-butoxycarbonyl-*O*-allyl-L-serine **3.35** (0.982 g, 84%) as a colourless oil; $[\alpha]_D^{20}$ +16 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D$ +9.2 (*c*, 1.1 in CH_2Cl_2)][18]; υ_{max} (neat) 1712 cm^{-1} (C=O); δ_H (400 MHz, $CDCl_3$)[18] 1.44 (9H, s, 3 x C$\underline{H}_3$), 3.67 (1H, dd, *J* 9.4 Hz, *J* 3.1 Hz, CHC$\underline{H}$H'), 3.89 (1H, d, *J* 8.2 Hz, CHCH$\underline{H}$'), 4.00 (2H, d, *J* 5.5 Hz, OC$\underline{H}_2$CH=CH_2), 4.44 (1H, br. s., C$\underline{H}CH_2$), 5.18 (1H, d, *J* 10.6 Hz, OCH_2CH=C$\underline{H}_Z H_E$), 5.25 (1H, d, *J* 17.2 Hz, OCH_2CH=$CH_Z\underline{H}_E$), 5.41 (1H, d, *J* 7.4 Hz, N$\underline{H}$), 5.76 - 5.93 (1H, m, $OCH_2C\underline{H}$=CH_2); δ_C (100.5 MHz, $CDCl_3$)[18] 28.3 (q, 3 x $\underline{C}H_3$), 53.8 (d, $\underline{C}HCH_2$), 69.6 (t, $CH\underline{C}H_2$), 72.4 (t, $O\underline{C}H_2CH$=CH_2), 80.3 (s, $\underline{C}(CH_3)_3$), 117.7 (t, CH_2CH=$\underline{C}H_2$), 133.9 (d, $CH_2\underline{C}H$=CH_2), 155.7 (s, O$\underline{C}$(O)NH), 175.1 (s, $\underline{C}$(O)OH); HRMS (ESI-TOF): calcd. for $C_{11}H_{19}NO_5Na^+$: 268.1155. Found: 268.1160 (MNa^+).

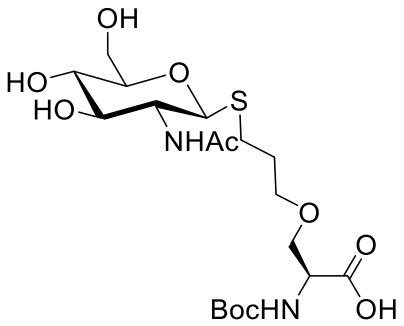
6-*O*-(2-[1-Thio-2-deoxy-2-acetamido-β-D-glucopyranoside]-ethyl)-D-galactopyranoside 3.36



1-Thio-2-deoxy-2-acetamido-β-D-glucopyranoside **3.5** (104 mg, 0.44 mmol) and 6-*O*-allyl-D-galactopyranoside **3.36** (88 mg, 0.4 mmol) were dissolved in water (4 mL). 2,2-Dimethoxy-2-phenylacetophenone (10.3 mg, 0.04 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 2 h. The reaction mixture

was then concentrated *in vacuo* and the residue purified by flash column chromatography ($CHCl_3$:MeOH 5:1) to afford 6-*O*-(2-[1-thio-2-deoxy-2-acetamido-β-D-glucopyranoside]-ethyl)-D-galactopyranoside **3.36** (133.1 mg, 73 %) as a colourless solid (α:β, 1:1.7); υ_{max} (neat) 3265 cm^{-1} (br. s, OH), 1637 cm^{-1} (C=O); δ_H (400 MHz, D_2O) 1.82 - 1.95 (5.4H, m, 2 x $SCH_2C\underline{H}_2CH_2O$), 2.02 (8.1H, s, 2 x $C(O)C\underline{H}_3$), 2.62 - 2.86 (5.4H, m, 2 x $SC\underline{H}_2CH_2CH_2O$), 3.41 - 3.96 (20.4H, m, H-2_1α, H-3_1α, H-4_1α, H-6_1α, H-6'$_1$α, H-2_1β, H-3_1β, H-4_1β, H-5_1β, H-6_1β, H-6'$_1$β, H-2_2, H-3_2, H-4_2, H-5_2, H-6_2, H-6'$_2$ & 2 x $SCH_2CH_2C\underline{H}_2O$), 4.19 (1H, t, *J* 5.5 Hz, H-5_1α), 4.55 (1H, d, $J_{1,2}$ 7.8 Hz, H-1_1β), 4.60 (1.6H, d, $J_{1,2}$ 10.6 Hz, 2 x H-1_2), 5.23 (0.6H, d, $J_{1,2}$ 3.5 Hz, H-1_1α); δ_C (100.5 MHz, D_2O) 22.1 (q, $C(O)\underline{C}H_3$), 26.9 (t, $S\underline{C}H_2CH_2CH_2O$), 28.9 (t, $SCH_2\underline{C}H_2CH_2O$), 54.7 (d, C-$2_2$), 60.8 (t, C-$6_2$), 68.2, 68.5, 68.9, 69.0, 69.4, 69.5, 69.6, 69.7, 69.8, 71.7, 72.7, 73.2, 75.1, 79.8 (d, C-5_2), 84.3, 84.4 (2 x d, 2 x C-1_2), 92.2 (d, C-1_1α), 96.3 (d, C-1_1β), 174.3 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{17}H_{32}NO_{11}S^+$: 458.1691. Found: 458.1689 (MH^+).

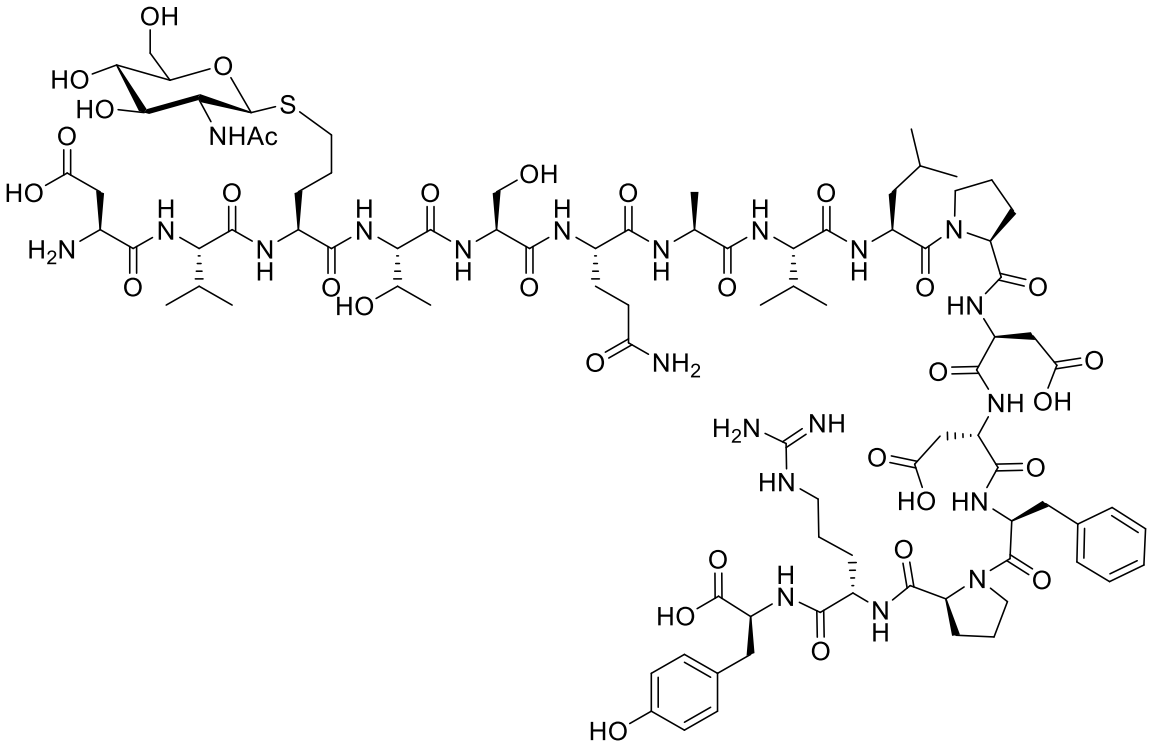
***N-tert*-Butoxycarbonyl-*O*-(3-[1-thio-2-deoxy-2-acetamido-β-D-glucopyranoside]-propyl)-L-serine 3.38**



1-Thio-2-deoxy-2-acetamido-β-D-glucopyranoside **3.5** (52 mg, 0.22 mmol) and *N-tert*-butoxycarbonyl-*O*-allyl-L-serine **3.35** (49 mg, 0.2 mmol) were dissolved in a 1:1 mixture of water and MeOH (2 mL). 2,2-Dimethoxy-2-phenylacetophenone (5.1 mg, 0.02 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor

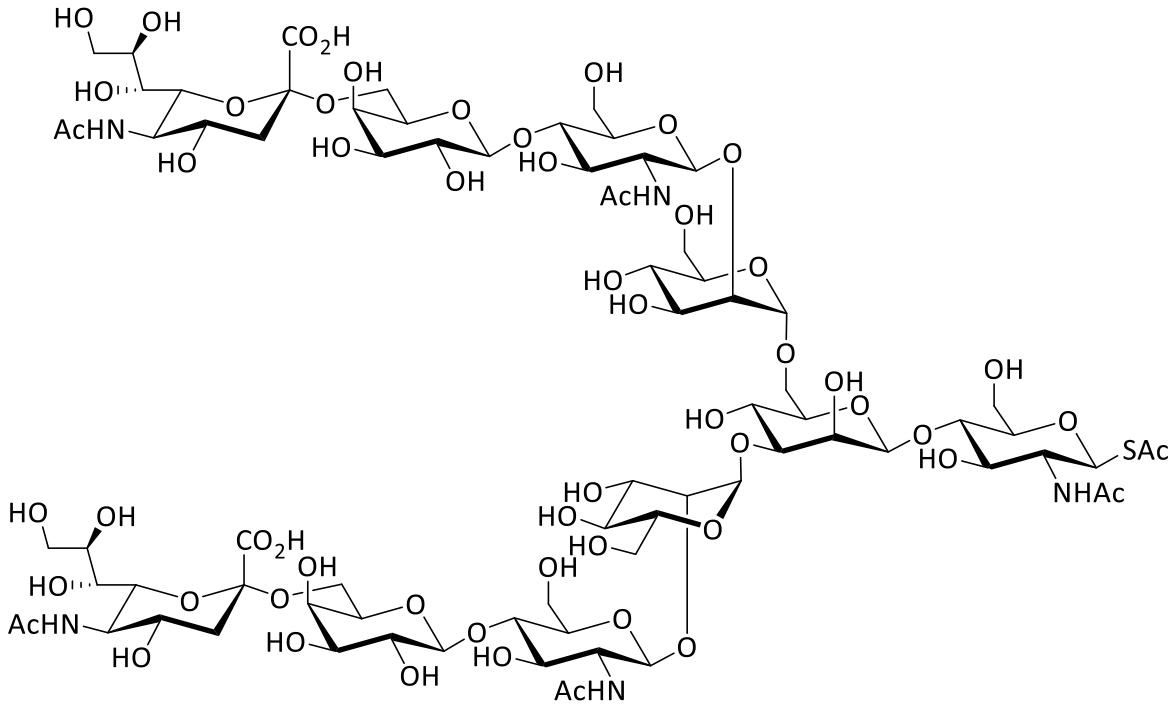
(wavelength: 254 nm) for 2 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by flash column chromatography ($CHCl_3$:MeOH 5:1) to afford *N-tert*-butoxycarbonyl-*O*-(2-[1-thio-2-deoxy-2-acetamido-β-D-glucopyranoside]-ethyl)-L-serine **3.38** (88 mg, 91%) as a colourless oil; $[\alpha]_D^{20}$ 0 (*c*, 1.0 in $CHCl_3$); υ_{max} (neat) 3306 cm^{-1} (OH), 1737 cm^{-1} (HOC=O), 1648 cm^{-1} (NHC=O); δ_H (400 MHz, CD_3OD) 1.44 (9H, s, $C(C\underline{H}_3)_3$), 1.78 - 1.89 (2H, m, $SCH_2C\underline{H}_2CH_2O$), 1.98 (3H, s, $NHC(O)C\underline{H}_3$), 2.64 - 2.84 (2H, m, $SC\underline{H}_2CH_2CH_2O$), 3.26-3.37 (2H, m, H-4 & H-5), 3.45 (1H, t, *J* 8.4 Hz, H-3), 3.50 - 3.59 (2H, m, $SCH_2CH_2C\underline{H}_2O$), 3.63 - 3.81 (4H, m, H-2, H-6 & $CHC\underline{H}_2$), 3.87 (1H, d, $J_{6,6'}$ 12.1 Hz, H-6'), 4.21 (1H, br. s, $C\underline{H}CH_2$) 4.48 (1H, d, *J* 10.2 Hz, H-1); δ_C (100.5 MHz, CD_3OD) 21.6 (q, $NHC(O)\underline{C}H_3$), 26.3 (t, $S\underline{C}H_2CH_2CH_2O$), 27.3 (q, $C(\underline{C}H_3)_3$), 29.4 (t, $SCH_2\underline{C}H_2CH_2O$), 54.5 (d, $\underline{C}HCH_2$), 54.8 (d, C-2), 61.5 (t, C-6), 69.3 (t, $SCH_2CH_2\underline{C}H_2O$), 70.4 (t, $CH\underline{C}H_2$), 70.6 (d, C-4), 76.0 (d, C-3), 80.6 (d, C-5), 84.4 (d, C-1); HRMS (ESI-TOF): calcd. for $C_{19}H_{35}N_2O_{10}S^+$: 483.2007. Found: 483.2012 (MH^+).

Neoglycopeptide 3.40



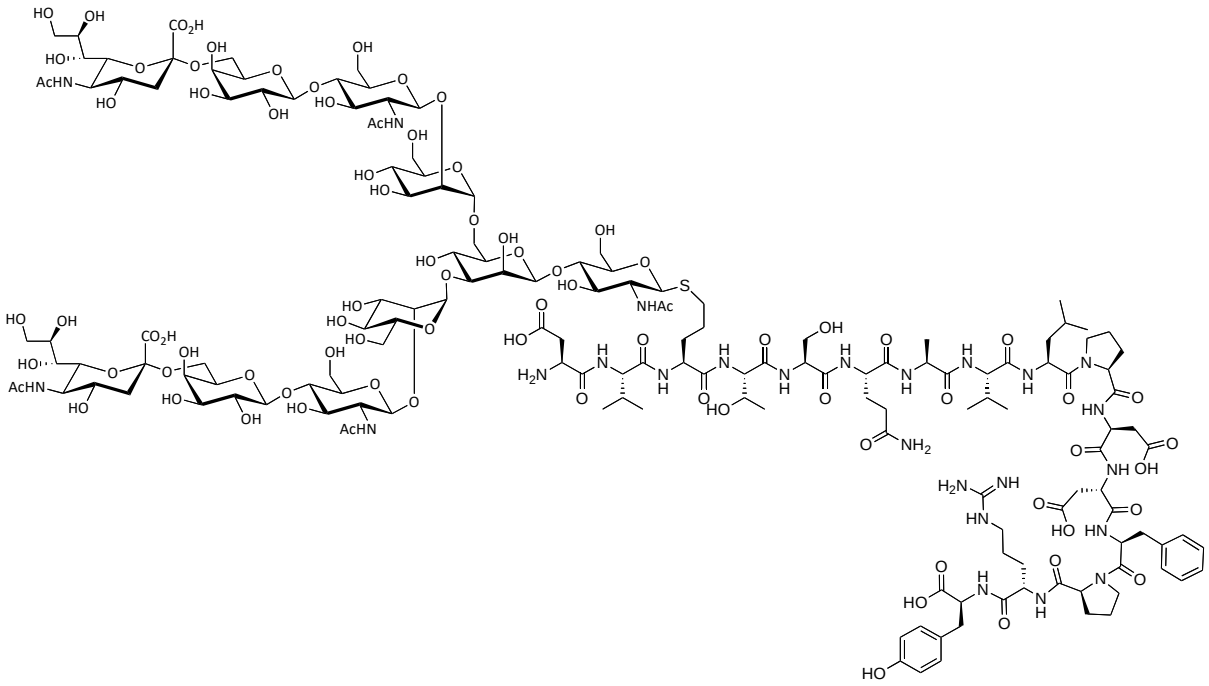
Peptide **3.39** (1.0 mg, 0.55 µmol) and 1-thio-2-acetamido-2-deoxy-β-D-glucopyranose **3.5** (6.5 mg, 27 µmol) were dissolved in DMF (100 µL). 2,2-Dimethoxy-2-phenylacetophenone (0.7 mg, 2.8 µmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 2 h. The reaction was analysed by LCMS (column: Jupiter 5u C18 (300 Å) column (Phenomenex); gradient: 5% MeCN in water for 15 min, followed by an increase to 75% MeCN over 30 min; column oven: 40 °C; flow rate: 0.5 mL/min; detection: UV 210 nm) to reveal the presence of neoglycopeptide **3.40** (>95%); t_R = 33.2 min; HRMS (ESI-TOF): calcd. for $C_{90}H_{138}N_{21}O_{32}S^+$: 2056.9532. Found: 2056.9590 (MH^+).

Sialoglycan thioacetate 3.42



Sialoglycan **3.41** (5.1 mg, 2.5 μmol) was dissolved in D_2O (100 μL). Triethylamine (35.0 μL, 250 μmol) was added and the reaction mixture was cooled to 0 °C and stirred. A solution of 2-chloro-1,3-dimethylimidazolinium chloride **2.1** (12.6 mg, 75 μmol) in D_2O (50 μL) was then added dropwise. The reaction mixture was stirred at 0 °C for 2 h. A solution of thioacetic acid (26.4 μL, 375 μmol) in MeCN (75 μL) was added dropwise to the reaction mixture. The reaction mixture was stirred for 1 h and then purified by HPLC (t_R = 14.8 min; column: XBridge® Glycan BEH Amide 3.5μm (130 Å) column; gradient: 10% water in MeCN for 7 min, followed by an increase to 40% water over 5 min, followed by 40% water in MeCN for 5 min; column oven: 40 °C; flow rate: 1 mL/min; detection: UV 210 nm) to afford sialoglycan thioacetate **3.42** (3.1 mg, 60%) as a white powder; δ_H (400 MHz, D_2O) 2.35 (3H, s, SC(O)$\underline{H}_3$), 5.17 (1H, d, $J_{1,2}$ 11.0 Hz, H-1_a); δ_C (100.5 MHz, D_2O) 30.0 (q, SC(O)$\underline{C}H_3$), 80.8 (d, C-1_a); HRMS (ESI-TOF): calcd. for $C_{78}H_{128}N_5O_{57}S^+$: 2078.6986. Found: 2078.7007 (MH^+).

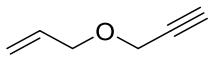
Sialoglycan neoglycopeptide 3.44



Sialoglycan thioacetate **3.42** (1.5 mg, 0.75 μmol) was dissolved in aqueous NaOH (1 M, 5 mL) and stirred for 2 h. Amberlite® IR-120 (H^+) was added until pH 7 was reached, and the reaction mixture was then filtered and concentrated *in vacuo*. The resulting colourless solid was dissolved in water (40 μL). Peptide **3.39** (0.5 mg, 0.28 μmol) was dissolved in DMF (100 μL). 10 μL of the resulting solution was diluted with water (150 μL). 5 μL of the diluted peptide solution was added to the sialoglycan solution. 2,2-Dimethoxy-2-phenylacetophenone (1 mg, 3.9 μmol) was dissolved in DMF (200 μL), and 1 μL of this solution was added to the reaction mixture. The mixture was then stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h, and then analysed by LCMS (column: Jupiter 5u C18 (300 Å) column (Phenomenex); gradient: 5% MeCN in water for 15 min, followed by an increase to 75% MeCN over 30 min; column oven: 40 °C; flow rate: 0.5 mL/min; detection: UV 235 nm) to reveal the presence of sialoglycan neoglycopeptide **3.44** (>95%); t_R = 32.1 min; HRMS (ESI-TOF): calcd. for $C_{158}H_{248}N_{25}O_{83}S^+$: 3855.5669. Found: 3855.5333 (MH^+).

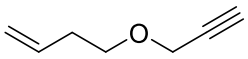
Experimental section for Chapter 4

3-Allyloxy-1-propyne 4.3[19]



NaH (60% suspension in mineral oil, 9 g, 225 mmol) was washed with diethyl ether (100 mL x 4) and dissolved in DMF/Et_2O (200 mL, 10:1) at 0 °C under an atmosphere of nitrogen. Propargyl alcohol (5.76 mL, 86 mmol) was added dropwise and the reaction was stirred for 1 h. Allyl bromide (7.38 mL, 99 mmol) was dissolved in Et_2O (45 mL) and the resulting solution was added dropwise to the reaction. The reaction was then stirred at 0 °C for 5 h, then quenched with brine (500 mL). The organic layer was washed with water (500 mL), dried ($MgSO_4$), filtered through silica, and carefully concentrated *in vacuo* (700 mbar, 40 °C) to afford 3-allylloxy-1-propyne **4.3** (7.24 g, 87%) as a pale yellow oil; δ_H (400 MHz, $CDCl_3$)[20] 2.40 (1H, d, *J* 1.2 Hz, $CH_2C{\equiv}C\underline{H}$), 4.03 (2H, d, *J* 5.9 Hz, $CH_2{=}CHC\underline{H}_2$), 4.11 (2H, d, *J* 2.3 Hz, $C\underline{H}_2C{\equiv}CH$), 5.18 (1H, d, *J* 10.2 Hz, $CH_E\underline{H}_Z{=}CH$), 5.27 (1H, d, *J* 17.2 Hz, $C\underline{H}_EH_Z{=}CH$), 5.86 (1H, ddt, J_E 16.9 Hz, J_Z 11.1 Hz, 2 x *J* 5.6 Hz, $CH_2{=}C\underline{H}CH_2$); δ_C (100.5 MHz, $CDCl_3$)[20] 57.0 (t, $\underline{C}H_2C{\equiv}CH$), 70.6 (t, $CH_2{=}CH\underline{C}H_2$), 74.3 (d, $CH_2C{\equiv}\underline{C}H$), 79.6 (s, $CH_2\underline{C}{\equiv}CH$), 117.9 (t, $\underline{C}H_2{=}CHCH_2$), 133.8 (d, $CH_2{=}\underline{C}HCH_2$).

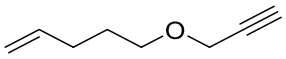
4-Prop-2-ynyloxy-but-1-ene 4.4[21]



Potassium hydroxide (5.05 g, 90 mmol) was suspended in DMSO (60 mL) at 0 °C. 3-Buten-1-ol (2.58 mL, 30 mmol) was added and the reaction mixture was stirred for 10 min. Propargyl bromide (2.27 mL, 30 mmol) was added and the reaction was stirred at RT for 3 h, then diluted with water (120 mL) and extracted with Et_2O (3 x 60 mL). The combined organic extracts were washed with water (4 x 50 mL), washed with brine (50 mL), dried ($MgSO_4$), filtered, and

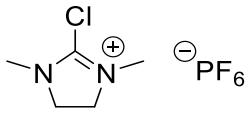
concentrated *in vacuo*. The residue was purified by flash column chromatography (petrol:Et_2O 98:2) to afford 4-prop-2-ynyloxy-but-1-ene **4.4** (1.31 g, 35%) as a pale yellow liquid; δ_H (400 MHz, $CDCl_3$) 2.32 - 2.44 (3H, m, C≡C<u>H</u> & C<u>H</u>$_2$CH=CH_2), 3.59 (2H, t, *J* 6.7 Hz, $CH_2C\underline{H}_2O$), 4.15 (2H, s, C<u>H</u>$_2$C≡CH), 5.06 (1H, d, *J* 10.6 Hz, $CH_E\underline{H}_Z$=CH), 5.12 (1H, d, *J* 17.2 Hz, C<u>H</u>$_E$H$_Z$=CH), 5.76 - 5.89 (1H, m, CH_2=C<u>H</u>CH_2); δ_C (100.5 MHz, $CDCl_3$) 33.9 (t, <u>C</u>H_2CH=CH_2), 58.0 (t, <u>C</u>H_2C≡CH), 69.3 (t, $CH_2\underline{C}H_2O$), 74.2 (d, C≡<u>C</u>H), 79.8 (s, $CH_2\underline{C}$≡CH), 116.5 (t, CH=<u>C</u>H_2), 134.9 (d, <u>C</u>H=CH_2).

5-Prop-2-ynyloxy-pent-1-ene 4.5[21]



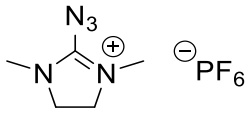
Potassium hydroxide (5.05 g, 90 mmol) was suspended in DMSO (60 mL) at 0 °C. 4-Penten-1-ol (2.58 mL, 30 mmol) was added and the mixture was stirred for 10 min. Propargyl bromide (2.27 mL, 30 mmol) was added and the mixture was stirred at RT for 3 h, then diluted with water (120 mL) and extracted with Et_2O (3 x 60 mL). The combined organic extracts were washed with water (4 x 50 mL), washed with brine (50 mL), dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The residue was purified by flash column chromatography (petrol:Et_2O 98:2) to afford 5-prop-2-ynyloxy-pent-1-ene **78** (1.80 g, 48%) as a pale yellow liquid; δ_H (400 MHz, $CDCl_3$) 1.62 - 1.73 (2H, m, $CH_2C\underline{H}_2CH_2$), 2.11 (2H, q, *J* 7.0 Hz, CH_2=CHC<u>H</u>$_2$), 2.39 (1H, t, *J* 2.3 Hz, C≡C<u>H</u>), 3.50 (2H, t, *J* 6.5 Hz, C<u>H</u>$_2$O), 4.11 (2H, d, *J* 2.7 Hz, C<u>H</u>$_2$≡CH), 4.95 (1H, dd, J_Z 10.2 Hz, J_{gem} 2.0 Hz, $CH_E\underline{H}_Z$=CH), 5.01 (1H, dd, J_E 17.0 Hz, J_{gem} 1.8 Hz, C<u>H</u>$_E$H$_Z$=CH), 5.79 (1H, ddt, J_E 17.1 Hz, J_Z 10.3 Hz, 2 x *J* 6.7 Hz, CH_2=C<u>H</u>CH_2).

2-Chloro-1,3-dimethylimidazolinium hexafluorophosphate 4.6[22]



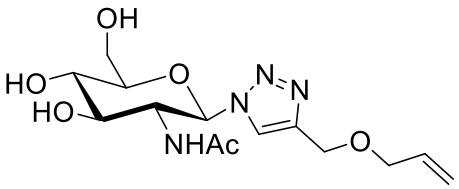
2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (20.00 g, 0.17 mol) and sodium hexafluophosphate (20.00 g, 0.17 mol) were dissolved in anhydrous MeCN under an atmosphere of nitrogen. The mixture was stirred for 0.5 h, then filtered through Celite® and concentrated *in vacuo*. The solid was precipitated ($MeCN/Et_2O$) to afford 2-chloro-1,3-dimethylimidazolinium **4.6** (26.03 g, 55%) as a white solid; m.p. 204-205 °C (MeCN) [lit. 210 °C (MeCN)][22]; δ_H (400 MHz, CD_3CN)[22] 3.13 (6H, s, 2 x CH_3), 3.95 (4H, s, 2 x CH_2); δ_C (100.5 MHz, CD_3CN)[22] 34.1 , 49.8; δ_F (376 MHz, CD_3CN)[22] -73.0 (d, *J* 707 Hz, P$\underline{F}_6$); δ_P (162 MHz, CD_3CN) -144.7 (sept., *J* 707 Hz, $\underline{P}F_6$).

2-Azido-1,3-dimethylimidazolinium hexafluorophosphate 4.7[22]



2-Chloro-1,3-dimethylimidazolinium chloride **4.6** (26.03 g, 0.093 mol) was dissolved in anhydrous MeCN (250 mL) at 0 °C under an atmosphere of nitrogen. Sodium azide (8.5 g, 0.13 mol) was added to the reaction and the reaction was stirred for 3 h. The reaction was then filtered through Celite® washed with MeCN, concentrated *in vacuo*, and the solid was precipitated ($MeCN/Et_2O$) to afford 2-azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (22.48 g, 85%) as a white solid; m.p. 206-207 °C (MeCN) [lit. 200-205 °C (MeCN)][22]; δ_H (400 MHz, CD_3CN)[22] 3.07 (6H, s, 2 x CH_3), 3.81 (4H, s, 2 x CH_2); δ_C (100.5 MHz, CD_3CN) 32.8, 48.9, 110.0; δ_F (376 MHz, CD_3CN) -73.1 (d, *J* 707 Hz, P$\underline{F}_6$); δ_P (162 MHz, CD_3CN) -144.7 (sept., *J* 707 Hz, $\underline{P}F_6$).

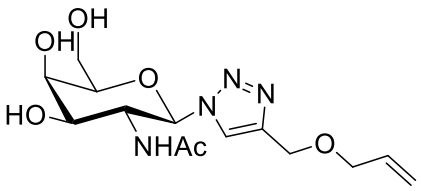
3-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene 4.9[22]



N-Acetyl-D-glucosamine **4.8** (500 mg, 2.26 mmol) and triethylamine (1.55 mL, 11.3 mmol) were dissolved in D_2O/MeCN (4:1, 10 mL) and the reaction mixture was cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (1.925 g, 6.75 mmol) was added and the reaction mixture was stirred for 1 h. The reaction mixture was then acidified to pH 2 by addition of HCl (1 M, 10 mL), stirred for 10 min, and then neutralised by the addition of NaHCO3 (sat. aqueous soln., 20 mL). The mixture was concentrated *in vacuo*. The residue was dissolved in *tert*-butanol/water (2:1, 25 mL) and 3-allyloxy-1-propyne **4.3** (0.45 mL, 4.0 mmol), $CuSO_4.5H_2O$ (112 mg, 0.45 mmol), and sodium ascorbate (178 mg, 0.9 mmol) were added. The reaction mixture was then heated to 50 °C and stirred for 16 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (50 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (50 mL), and the aqueous layer was washed with CH_2Cl_2 (2 x 100 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 8:1) to afford 3-((1'-(2"-acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.9** (615 mg, 79%) as a white solid; m.p. 95-100 °C (CH_2Cl_2); $[\alpha]_D^{20}$ -5.9 (*c*, 1.0 in MeOH); υ_{max} (neat) 3410 cm^{-1} (OH), 1651 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.80 (3H, s, NHC(O)C$\underline{H}_3$), 3.66 - 3.85 (4 H, m, H-3, H-4, H-5 & H-6), 3.92 (1H, dd, $J_{5,6'}$ 1.6 Hz, $J_{6,6'}$ 12.5 Hz, H-6'), 4.04 (2H, d, *J* 6.3 Hz, CH_2=CHC$\underline{H}_2$), 4.25 (1H,

t, *J* 10.0 Hz, H-2), 4.66 (2H, s, NCH=CC$\underline{H}_2$), 5.26 (1H, d, J_Z 10.6 Hz, $CH_E\underline{H}_Z$=CHCH$_2$), 5.32 (1H, d, J_E 17.2 Hz, C$\underline{H}_E$H$_Z$=CHCH$_2$), 5.84 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 5.92 (1H, ddt, J_E 17.2 Hz, J_Z 10.6, 2 x *J* 5.9 Hz, CH$_2$=C$\underline{H}$CH$_2$), 8.22 (1H, s, NC$\underline{H}$=C); δ_C (100.5 MHz, D$_2$O) 21.5 (q, NHC(O)$\underline{C}$H$_3$), 55.3 (d, C-2), 60.3 (t, C-6), 61.8 (t, NCH=C$\underline{C}$H$_2$), 69.2 (d, C-4), 70.8 (t, CH$_2$=CH$\underline{C}$H$_2$), 73.4 (d, C-3), 78.8 (d, C-5), 86.3 (d, C-1), 118.7 (t, $\underline{C}$H$_2$=CHCH$_2$), 123.7 (d, N$\underline{C}$H=C), 133.2 (d, CH$_2$=$\underline{C}$CH$_2$), 144.0 (s, NCH=$\underline{C}$CH$_2$), 174.0 (s, NH$\underline{C}$(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{14}H_{23}N_4O_6^+$: 343.1612. Found: 343.1611 (MH^+).

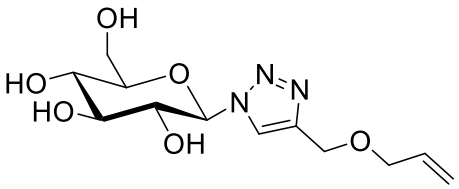
3-((1'-(2"-Acetamido-2"-deoxy-β-D-galactopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene 4.10[22]



N-Acetyl-D-galactosamine (111 mg, 0.5 mmol) and triethylamine (0.35 mL, 2.5 mmol) were dissolved in D$_2$O/MeCN (4:1, 2 mL) and the reaction mixture was cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (428 mg, 1.5 mmol) was added and the reaction mixture was stirred for 1 h. The reaction mixture was then acidified to pH 2 by addition of aqueous HCl (1 M, 2.5 mL), stirred for 10 min, and then neutralised by the addition of NaHCO3 (sat. aqueous soln., 5 mL). The mixture was concentrated *in vacuo*. The residue was dissolved in *tert*-butanol/water (2:1, 6 mL) and 3-allyloxy-1-propyne **4.3** (0.11 mL, 2.0 mmol), CuSO$_4$.5H$_2$O (25 mg, 0.5 mmol), and sodium ascorbate (40 mg, 0.2 mmol) were added. The reaction mixture was then heated to 50 °C and stirred for 16 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (20 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (10 mL), and the aqueous layer was washed with CH$_2$Cl$_2$ (3 x 20 mL), filtered through a column of Amberlite® IR-120 (Na^+

form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 8:1) to afford 3-((1'-(2"-acetamido-2"-deoxy-β-D-galactopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.10** (75.3 mg, 44%) as a white solid; $[\alpha]_D^{20}$ +6 (*c*, 1.0 in MeOH); υ_{max} (neat) 3290 cm^{-1} (OH), 1652 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.80 (3H, s, NHC(O)C<u>H</u>$_3$), 3.79 - 3.84 (2H, m, H-6 & H-6'), 3.95 - 4.02 (2H, m, H-3 & H-5), 4.05 (2H, d, *J* 5.9 Hz, CH_2=CHC<u>H</u>$_2$), 4.10 (1H, d, $J_{3,4}$ 3.1 Hz, H-4), 4.42 (1H, t, *J* 10.2 Hz, H-2), 4.67 (2H, s, CC<u>H</u>$_2$O), 5.27 (1H, dd, J_Z 10.6 Hz, J_{gem} 0.8 Hz, C<u>H</u>$_Z$H$_E$=CH), 5.32 (1H, dd, J_E 17.4 Hz, J_{gem} 1.4 Hz, CH$_Z$<u>H</u>$_E$=CH), 5.78 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 5.92 (1H, ddt, J_E 17.1 Hz, J_Z 10.7 Hz, 2 x *J* 5.9 Hz, CH_2=C<u>H</u>), 8.27 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 21.6 (q, <u>C</u>H$_3$), 52.0 (d, C-2), 60.8 (t, C-6), 61.8 (t, C<u>C</u>H$_2$O), 67.6 (d, C-4), 70.6 (d, C-3), 70.8 (t, CH_2=CH<u>C</u>H$_2$), 78.3 (d, C-5), 86.9 (d, C-1), 118.8 (t, <u>C</u>H$_2$=CH), 123.6 (d, N<u>C</u>H=C), 133.3 (d, CH_2=<u>C</u>H), 144.1 (s, NCH=<u>C</u>CH$_2$), 174.2 (s, NH<u>C</u>(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{14}H_{23}N_4O_6^+$: 343.1612. Found: 343.1613 (MH^+).

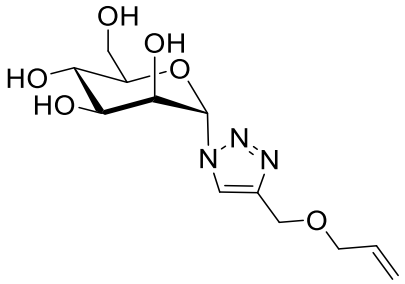
3-((1'-β-D-Glucopyranosyl-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene 4.11[22]



Glucose (90.1 mg, 0.5 mmol) and triethylamine (0.35 mL, 2.5 mmol) were dissolved in D_2O/MeCN (4:1, 2 mL) and the reaction mixture was cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (428 g, 1.5 mmol) was added and the reaction mixture was stirred for 1 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in *tert*-butanol/water (2:1, 6 mL). 3-Allyloxy-1-propyne **4.3** (0.11 mL, 1.0 mmol), $CuSO_4.5H_2O$ (25.0 mg, 0.1 mmol), and sodium ascorbate (39.6 mg, 0.2 mmol)

were added. The reaction mixture was then heated to 50 °C and stirred for 16 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (25 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (20 mL), and the aqueous layer was washed with CH_2Cl_2 (3 x 20 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 8:1) to afford 3-((1'-β-D-glucopyranosyl-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.11** (100 mg, 66%) as a colourless oil; $[\alpha]_D^{20}$ -3 (*c*, 1.0 in MeOH); υ_{max} 3334 cm^{-1} (OH); δ_H (400 MHz, D_2O) 3.58 - 3.64 (1H, m, H-4), 3.67 - 3.80 (3H, m, H-3, H-5 & H-6'), 3.90 (1H, d, $J_{6,6'}$ 10.6 Hz, H-6'), 3.99 (1H, t, *J* 9.2 Hz, H-2), 4.10 (2H, d, *J* 5.9 Hz, CHC$\underline{H}_2$), 4.70 (2H, s, CC$\underline{H}_2$O), 5.27 (1H, d, J_Z 10.6 Hz, C$\underline{H}_Z$H$_E$=CH), 5.33 (1H, d, J_E 17.2 Hz, CH$_Z\underline{H}_E$=CH), 5.75 (1H, d, $J_{1,2}$ 9.4 Hz, H-1), 5.94 (1H, ddt, J_E 17.0 Hz, J_Z 10.6 Hz, 2 x *J* 6.2 Hz, CH_2=C$\underline{H}CH_2$), 8.25 (1H, s, NC$\underline{H}$=C); δ_C (100.5 MHz, D_2O) 60.3 (t, C-6), 62.0 (t, C$\underline{C}H_2$O), 68.9 (d, C-4), 71.1 (t, CH_2=CH$\underline{C}H_2$), 72.2 (d, C-2), 75.8 (d, C-3), 78.8 (d, C-5), 87.4 (d, C-1), 118.8 (t, $\underline{C}H_2$=CH), 124.3 (d, N$\underline{C}$H=C), 133.3 (d, CH_2=$\underline{C}$H), 144.2 (s, NCH=$\underline{C}CH_2$); HRMS (ESI-TOF): calcd. for $C_{12}H_{20}N_3O_6^+$: 302.1347. Found: 302.1348 (MH^+).

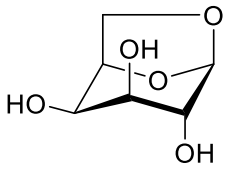
3-((1'-α-D-Mannopyranosyl-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene 4.12[22]



Mannose (90.1 mg, 0.5 mmol) and triethylamine (0.35 mL, 2.5 mmol) were dissolved in D_2O/MeCN (4:1, 2 mL) and the reaction mixture was cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (428 g, 1.5 mmol) was added and the reaction mixture was stirred for 1 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in *tert*-butanol/water (2:1, 6 mL). 3-Allyloxy-1-propyne **4.3** (0.11 mL, 1.0 mmol), $CuSO_4.5H_2O$ (25.0 mg, 0.1 mmol), and sodium ascorbate (39.6 mg, 0.2 mmol) were added. The reaction mixture was then heated to 50 °C and stirred for 40 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (25 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (20 mL), and the aqueous layer was washed with CH_2Cl_2 (3 x 20 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 8:1) to afford 3-((1'-α-D-mannopyranosyl-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.12** (120 mg, 80%) as a colourless oil; $[\alpha]_D^{20}$ +42 (*c*, 1.0 in MeOH); υ_{max} 3353 cm^{-1} (OH); δ_H (400 MHz, D_2O) 3.27 - 3.32 (1H, m, H-5), 3.75 - 3.85 (3H, m, H-4, H-6 & H-6'), 4.07 - 4.16 (3H, m, H-3 & CHC$\underline{H}_2$O), 4.69 (2H, s, CC$\underline{H}_2$O), 4.79 (1H, t, *J* 3.1 Hz, H-2), 5.26 (1H, dd, J_Z 10.6 Hz, J_{gem} 0.8 Hz, C$\underline{H}_Z$H$_E$=CH), 5.32 (1H, dd, J_E 17.2 Hz, J_{gem} 1.6 Hz, CH$_Z\underline{H}_E$=CH), 5.93 (1H, ddt, J_E 17.0 Hz, J_Z 10.6 Hz, 2 x *J* 6.2 Hz, CH$_2$=C$\underline{H}$), 6.12 (1H, d, $J_{1,2}$ 2.3 Hz, H-1), 8.19 (1H, s, NC$\underline{H}$=C); δ_C (100.5 MHz, D_2O) 60.4 (t, C-6), 62.0 (t, C$\underline{C}H_2$O), 66.5 (d, C-4), 68.2 (d,

C-2), 70.5 (d, C-3), 71.1 (t, CH<u>C</u>H$_2$O), 76.1 (d, C-5), 86.7 (d, C-1), 118.7 (t, <u>C</u>H$_2$=CH), 124.7 (d, N<u>C</u>H=C), 133.4 (d, CH$_2$=<u>C</u>H), 144.3 (s, NCH=<u>C</u>CH$_2$); HRMS (ESI-TOF): calcd. for $C_{12}H_{20}N_3O_6^+$: 302.1347. Found: 302.1339 (MH^+).

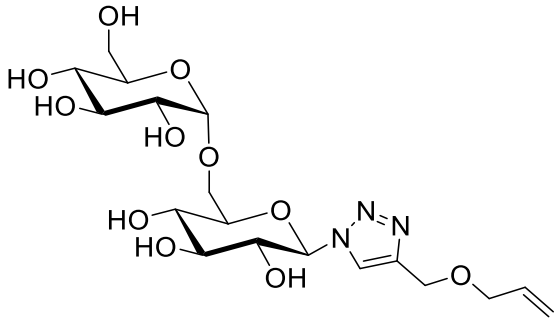
1,6-Anhydro-β-D-galactose 4.13[10]



Galactose (90.1 mg, 0.5 mmol) and triethylamine (0.35 mL, 2.5 mmol) were dissolved in D$_2$O/MeCN (4:1, 2 mL) and the reaction mixture was cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (428 g, 1.5 mmol) was added and the reaction mixture was stirred for 1 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in *tert*-butanol/water (2:1, 6 mL). 3-Allyloxy-1-propyne **4.3** (0.11 mL, 1.0 mmol), CuSO$_4$.5H$_2$O (25.0 mg, 0.1 mmol), and sodium ascorbate (39.6 mg, 0.2 mmol) were added. The reaction mixture was then heated to 50 °C and stirred for 16 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (25 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (20 mL), and the aqueous layer was washed with CH$_2$Cl$_2$ (3 x 20 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH$_2$Cl$_2$:MeOH 8:1) to afford 1,6-anhydro-β-D-galactose **4.13** (33.3 mg, 41%) as a colourless oil; δ_H (400 MHz, D$_2$O)[10] 3.63 (1H, dd, $J_{5,6}$ 5.7 Hz, $J_{6,6'}$ 7.2 Hz, H-6), 3.77 - 3.79 (1H, m, H-2), 3.90 - 3.93 (1H, m, H-3), 4.03 (1H, t, *J* 4.5 Hz, H-4), 4.29 (1H, d, $J_{6,6'}$ 7.4 Hz, H-6'), 4.48 (1H, t, *J* 4.5 Hz, H-

5), 5.38 (1H, s, H-1); δ_C (100.5 MHz, D_2O) 63.4 (t, C-6), 64.2 (d, C-4), 70.0 (d, C-3), 71.2 (d, C-2), 74.2 (d, C-5), 100.6 (d, C-1).

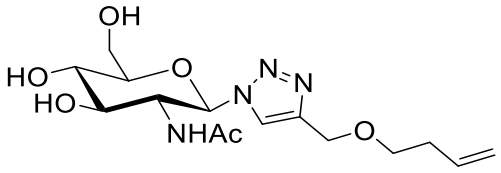
3-((1'-(α-D-Glucopyranosyl-(1→4)-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene 4.14[22]



Isomaltose (34.2 mg, 0.1 mmol) and triethylamine (0.07 mL, 0.5 mmol) were dissolved in D_2O/MeCN (4:1, 0.4 mL) and the reaction mixture was cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (85.5 g, 0.3 mmol) was added and the reaction mixture was stirred for 1 h. *tert*-Butanol (0.8 mL) was added to the reaction mixture, followed by 3-allyloxy-1-propyne **4.3** (0.02 mL, 0.2 mmol), $CuSO_4.5H_2O$ (5.0 mg, 0.02 mmol), and sodium ascorbate (7.9 mg, 0.04 mmol). The reaction was heated to 50 °C and stirred for 16 h, then concentrated *in vacuo*. The residue was dissolved in EtOH (5 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (5 mL), and the aqueous layer was washed with CH_2Cl_2 (3 x 10 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 4:1) to afford 3-((1'-(α-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.14** (34.7 mg, 75%) as a white solid; $[\alpha]_D^{20}$ +47 (*c*, 1.0 in MeOH); υ_{max} 3258 cm^{-1} (OH); δ_H (400 MHz, D_2O) 3.37 (1H, t, *J* 9.8 Hz, H-5$_b$), 3.51 (1H, dd, $J_{1,2}$ 3.9 Hz, $J_{2,3}$ 10.2 Hz, H-2$_b$),

3.59 - 3.76 (6H, m, H-3_a, H-5_a, H-3_b, H-4_b, H-6_b & H-6'$_b$), 3.76 - 4.00 (3H, m, H-4_a, H-6_a & H-6'$_a$), 4.06 (1H, t, *J* 9.2 Hz, H-2_a), 4.11 (2H, d, *J* 5.9 Hz, CH_2=CHC$\underline{H}_2$), 4.71 (2H, s, CC$\underline{H}_2$O), 4.91 (1H, d, $J_{1,2}$ 3.9 Hz, H-1_b), 5.28 (1H, d, J_Z 10.6 Hz, C$\underline{H}_Z$$H_E$=CH), 5.34 (1H, dd, J_E 17.2 Hz, J_{gem} 1.6 Hz, $CH_Z$$\underline{H}_E$=CH), 5.77 (1H, d, $J_{1,2}$ 9.4 Hz, H-1_a), 5.95 (1H, ddt, J_E 17.0 Hz, J_Z 10.6 Hz, 2 x *J* 6.2 Hz, CH_2=C$\underline{H}$), 8.28 (1H, s, NC$\underline{H}$=C); δ_C (100.5 MHz, D_2O) 60.0 (t, C-6_b), 61.9 (t, C$\underline{C}H_2$O), 65.9 (t, C-6_a), 69.0 (d, C-5_a), 69.2 (d, C-5_b), 71.1 (t, CH_2=CH$\underline{C}H_2$), 71.3 (d, C-2_b), 71.6 (d, C-4_b), 72.0 (d, C-2_a), 72.9 (d, C-3_b), 76.0 (d, C-3_a), 77.5 (d, C-4_a), 87.3 (s, C-1_a), 98.0 (s, C-1_b), 118.8 (t, $\underline{C}H_2$=C), 124.5 (d, N$\underline{C}$H=C), 133.3 (s, CH=$\underline{C}CH_2$); HRMS (ESI-TOF): calcd. for $C_{18}H_{30}N_3O_{11}^+$: 464.1875. Found: 464.1878 (MH^+).

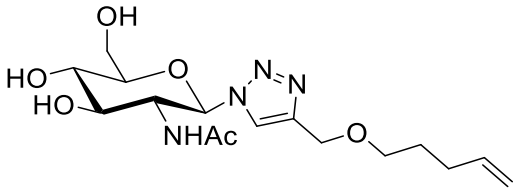
4-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-butene 4.15[22]



N-Acetyl-D-glucosamine **4.8** (221 mg, 1 mmol) and triethylamine (0.70 mL, 5 mmol) were dissolved in D_2O/MeCN (4:1, 4 mL) and the solution cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (855 g, 3 mmol) was added and the reaction mixture was stirred for 1 h. The reaction mixture was then acidified to pH 2 by the addition of aqueous HCl (1 M, 5 mL), stirred for 10 min, and then neutralised by addition of $NaHCO_3$ (sat. aqueous soln., 10 mL). The mixture was concentrated *in vacuo*. The residue was dissolved in *tert*-butanol/water (2:1, 10 mL) and 4-prop-2-ynyloxy-but-1-ene **4.4** (220 mg, 2 mmol), $CuSO_4.5H_2O$ (50 mg, 0.2 mmol), and sodium ascorbate (80 mg, 0.4 mmol) were added. The reaction mixture was then heated to 50 °C and stirred for 16 h. The mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (25 mL), filtered through Celite®,

and concentrated *in vacuo*. The residue was dissolved in water (25 mL), and the aqueous layer was washed with CH_2Cl_2 (2 x 50 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 8:1) to afford 4-((1'-(2"-acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-butene **4.15** (190 mg, 53%) as a white solid; υ_{max} (neat) 3279 cm^{-1} (OH), 1656 cm^{-1} (NHC=O); $[\alpha]_D^{20}$ -13.7 (*c*, 1.0 in MeOH); δ_H (400 MHz, D_2O) 1.80 (3H, s, C(O)C<u>H</u>$_3$), 2.32 (2H, q, *J* 6.5 Hz, CHC<u>H</u>$_2$CH$_2$O), 3.58 (2H, t, *J* 6.5 Hz, CHCH$_2$C<u>H</u>$_2$O), 3.65 - 3.85 (4H, m, H-3, H-4, H-5 & H-6), 3.93 (1H, d, $J_{6,6'}$ 12.5 Hz, H-6'), 4.25 (1H, t, *J* 9.8 Hz, H-2), 4.65 (2H, s, CC<u>H</u>$_2$O), 5.03 - 5.16 (2H, m, C<u>H</u>$_2$=CH), 5.74 - 5.89 (2H, m, H-1 & CH$_2$=C<u>H</u>), 8.21 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 21.6 (q, C(O)<u>C</u>H$_3$), 33.0 (t, CH<u>C</u>H$_2$CH$_2$O), 55.3 (d, C-2), 60.4 (t, C-6), 62.4 (t, C<u>C</u>H$_2$O), 69.1 (t, CHCH$_2$<u>C</u>H$_2$O), 69.2 (d, C-4), 73.5 (d, C-3), 78.9 (d, C-5), 86.3 (d, C-1), 116.6 (t, <u>C</u>H$_2$=CH), 123.7 (d, N<u>C</u>H=C), 135.3 (d, CH$_2$=<u>C</u>H), 144.1 (CH=<u>C</u>CH$_2$O), 174.0 (q, <u>C</u>(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{15}H_{25}N_4O_6^+$: 357.1769. Found: 357.1771 (MH^+).

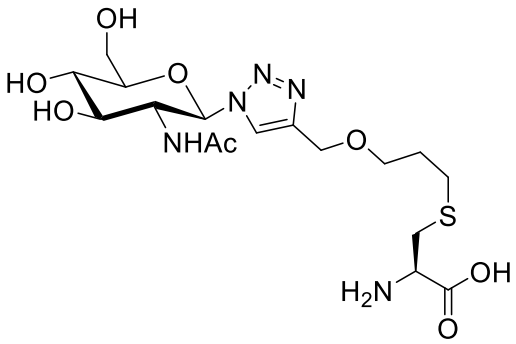
5-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-pentene 4.16[22]



N-Acetyl-D-glucosamine **4.8** (221 mg, 1 mmol) and triethylamine (0.70 mL, 5 mmol) were dissolved in D_2O/MeCN (4:1, 4 mL) and the solution cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (855 g, 3 mmol) was added and the mixture was stirred for 1 h. The reaction mixture was then acidified to pH 2 by addition of aqueous HCl

(1 M, 5 mL), stirred for 10 min, and then neutralised by the addition of $NaHCO_3$ (sat. aqueous soln., 10 mL). The mixture was concentrated *in vacuo*. The residue was dissolved in *tert*-butanol/water (2:1, 10 mL) and 5-prop-2-ynyloxy-pent-1-ene **4.5** (248 mg, 2 mmol), $CuSO_4.5H_2O$ (50 mg, 0.2 mmol), and sodium ascorbate (80 mg, 0.4 mmol) were added. The reaction mixture was then heated to 50 °C and stirred for 16 h. The reaction mixture was then concentrated *in vacuo* and the residue was dissolved in EtOH (25 mL), filtered through Celite®, and concentrated *in vacuo*. The residue was dissolved in water (25 mL), and the aqueous layer was washed with CH_2Cl_2 (2 x 50 mL), filtered through a column of Amberlite® IR-120 (Na^+ form, produced by treating the H^+ form with 1 M aqueous NaOH solution and then washing with water until pH 7), and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 8:1) to afford 5-((1'-(2"-acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-pentene **4.16** (329 mg, 89%) as a pale yellow oil; $[\alpha]_D^{20}$ -9 (*c*, 1.0 in MeOH); υ_{max} (neat) 3406 cm^{-1} (OH), 1656 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.65 (2H, quin, *J* 6.9 Hz, $CH_2C\underline{H}_2CH_2$), 1.80 (3H, s, $C(O)C\underline{H}_3$), 2.07 (2H, q, *J* 7.0 Hz, $CH_2{=}CHC\underline{H}_2$), 3.53 (2H, t, *J* 6.5 Hz, $CH_2C\underline{H}_2O$), 3.65 - 3.85 (4H, m, H-3, H-4, H-5 & H-6), 3.93 (1H, dd, $J_{5,6'}$ 2.0 Hz, $J_{6,6'}$ 12.1 Hz, H-6'), 4.26 (1H, t, *J* 10.0 Hz, H-2), 4.64 (2H, s, $CC\underline{H}_2O$), 4.94 - 5.06 (2H, m, $C\underline{H}_2{=}CHCH_2$), 5.78 - 5.91 (2H, m, H-1 & $CH_2{=}C\underline{H}CH_2$), 8.21 (1H, s, $NC\underline{H}{=}C$); δ_C (100.5 MHz, D_2O) 21.6 (q, $C(O)\underline{C}H_3$), 27.7 (t, $CH_2\underline{C}H_2CH_2$), 29.4 (t, $CH_2{=}CH\underline{C}H_2$), 55.2 (d, C-2), 60.4 (t, C-6), 62.3 (t, $C\underline{C}H_2O$), 69.2 (d, C-4), 69.4 (t, $CH_2\underline{C}H_2O$), 73.5 (d, C-3), 78.9 (d, C-5), 86.3 (d, C-1), 114.7 (t, $\underline{C}H_2{=}CH$), 123.7 (d, $N\underline{C}H{=}C$), 138.6 (d, $CH_2{=}\underline{C}HCH_2$), 144.2 (s, $CH{=}\underline{C}CH_2O$), 174.0 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{16}H_{27}N_4O_6^+$: 371.1925. Found: 371.1925 (MH^+).

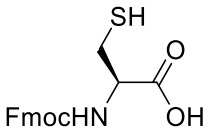
S-(3'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-L-cysteine 4.17



3-((1'-(2''-Acetamido-2''-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.9** (50 mg, 0.15 mmol) and *N*-[(9-fluorenylmethoxy)carbonyl]-L-cysteine **4.19** (60 mg, 0.17 mmol) were dissolved in a mixture of sodium acetate buffer (0.7 mL, pH 4.0, 0.6 M) and *tert*-butanol (1.3 mL). 2,2-Dimethoxy-2-phenylacetophenone (4.0 mg, 0.016 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 2 h. A further portion of *N*-[(9-fluorenylmethoxy)carbonyl]-L-cysteine **4.12** (60 mg, 0.17 mmol) was added, and the reaction mixture was stirred and irradiated for a further 2 h. The reaction mixture was then concentrated *in vacuo* and dissolved in anhydrous DMF (4.5 mL). Piperidine (0.5 mL, 6.8 mmol) was added, and the mixture stirred for 2 h. The mixture was concentrated *in vacuo* and the residue was purified by reverse phase flash column chromatography to afford *S*-(3'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)- L-cysteine **4.17** (58 mg, 74%) as a colourless oil; $[\alpha]_D^{20}$ -15 (*c*, 0.9 in water); υ_{max} 3245 cm^{-1} (OH), 1628 cm^{-1} (NC=O); δ_H (400 MHz, D_2O) 1.80 (3H, s, NHC(O)C$\underline{H}_3$), 1.86 (2H, quin, *J* 6.7 Hz, $CH_2C\underline{H}_2CH_2$), 2.63 (2H, t, *J* 6.8 Hz, $CH_2C\underline{H}_2S$), 2.98 (1H, dd, *J* 14.9 Hz, *J* 7.8 Hz, CHC$\underline{H}$H'S), 3.09 (2H, dd, *J* 14.9 Hz, *J* 4.3 Hz, CHCH$\underline{H}$'S), 3.61 (2H, t, *J* 6.1 Hz, $OC\underline{H}_2CH_2$), 3.65 - 3.85 (4H, m, H-3, H-4, H-5 & H-6), 3.86 - 3.97 (2H, m, C$\underline{H}CH_2$ & H-6'), 4.25 (1H, t, *J* 9.8 Hz, H-2), 4.64 (2H, s, $CC\underline{H}_2O$), 5.84 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 8.23 (1H, s, NC$\underline{H}$=C); δ_C (100.5 MHz, D_2O) 21.6

(q, NHC(O)<u>C</u>H$_3$), 27.9 (t, CH$_2$<u>C</u>H$_2$S), 28.3 (t, CH$_2$C<u>H</u>$_2$CH$_2$), 32.0 (t, CH<u>C</u>H$_2$S), 53.5 (d, <u>C</u>HCH$_2$), 55.3 (d, C-2), 60.3 (t, C-6), 62.4 (t, C<u>C</u>H$_2$O), 68.3 (t, O<u>C</u>H$_2$CH$_2$), 69.2 (d, C-4), 73.4 (d, C-3), 78.8 (d, C-5), 86.3 (d, C-1), 123.7 (d, N<u>C</u>H=C), 144.1 (s, CH=<u>C</u>CH$_2$), 172.7, 174.0 (2 x s, 2 x C=O); HRMS (ESI-TOF): calcd. for $C_{17}H_{30}N_5O_8S^+$: 464.1810. Found: 464.1816 (MH^+).

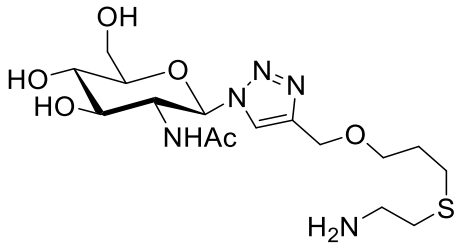
***N*-[(9-Fluorenylmethoxy)carbonyl]-L-cysteine 4.19[23]**



Trifluoroacetic acid (0.19 mL, 2.5 mmol) and triisopropylsilane (0.37 mL, 2.5 mmol) were added to a solution of *N*-[(9-fluorenylmethoxy)carbonyl]-*S*-trityl-L-cysteine **4.18** (293 mg, 0.5 mmol) in CH_2Cl_2 (5 mL) under an atmosphere of nitrogen. After 1 h, the reaction was concentrated *in vacuo*. The residue was redissolved in CH_2Cl_2 (20 mL) and concentrated *in vacuo*, and this process was repeated until a white solid was obtained in order to remove all trifluoroacetic acid. The solid was precipitated (CH_2Cl_2/petrol) to afford *N*-[(9-fluorenylmethoxy)carbonyl]-L-cysteine **4.19** (142 mg, 83%) was a white solid; $[\alpha]_D^{20}$ -25 (*c*, 2.0 in DMF) [lit. $[\alpha]_D^{20}$ -24.8 (*c*, 2.3 in DMF)][23]; υ_{max} (neat) 3310 cm^{-1} (NH) 1701 cm^{-1}, 1688 cm^{-1} (NC=O & HOC=O); δ_H (400 MHz, DMSO-d$_6$)[23] 2.65 - 2.76 (1H, m, C<u>H</u>H'SH), 2.82 - 2.91 (1H, m, CH<u>H</u>'SH), 4.09 (1H, td, *J* 8.4 Hz, *J* 4.3 Hz, C<u>H</u>CH$_2$SH), 4.21 (1H, t, *J* 7.0 Hz, C<u>H</u>CH$_2$O), 4.26 - 4.31 (2H, m, CHC<u>H</u>$_2$O), 7.25 - 7.35 (2H, m, 2 x Ar-H), 7.35 - 7.45 (2H, m, 2 x Ar-H), 7.66 (1H, d, *J* 8.6 Hz, NH), 7.71 (2H, d, *J* 7.4 Hz, 2 x Ar-H), 7.87 (2H, d, *J* 7.4 Hz, 2 x Ar-H); δ_C (100.5 MHz, DMSO-d$_6$)[23] 25.9 (t, <u>C</u>H$_2$SH), 47.1 (d, <u>C</u>HCH$_2$O), 57.0 (d, <u>C</u>HCH$_2$SH), 66.2 (t, CH<u>C</u>H$_2$O), 120.6 (d, Ar-C), 125.7 (d, Ar-C), 127.5 (D, Ar-C), 128.1 (d,

Ar-C), 141.2 (s, Ar-C), 144.2 (s, Ar-C), 144.2 (s, Ar-C), 156.5 (s, OC(O)NH), 172.3 (s, COOH); HRMS (ESI-TOF): calcd. for $C_{18}H_{17}NO_4SNa^+$: 366.0770. Found: 366.0756 (MNa^+).

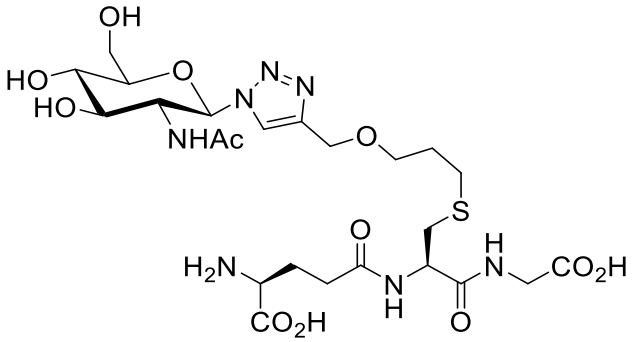
_S_-(3'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-1-thio-2-aminoethane 4.21



3-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.9** (41 mg, 0.12 mmol) and cysteamine hydrochloride (42 mg, 0.36 mmol) were dissolved in sodium acetate buffer (2 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (3.1 mg, 0.012 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 6 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by flash column chromatography (CH_2Cl_2:MeOH:NH_3 75:25:1) to afford *S*-(3'-((1''-(2'''-acetamido-2'''-deoxy-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-1-thio-2-aminoethane **4.21** (31.1 mg, 61%) as a colourless oil; υ_{max} (neat) 3246 cm^{-1} (broad, NH & OH), 1649 cm^{-1} (C=O); $[\alpha]_D^{20}$ -8 (*c*, 1.6 in MeOH); δ_H (400 MHz, D_2O) 1.80 (3H, s, C(O)C<u>H</u>$_3$), 1.85 (2H, quin, *J* 6.7 Hz, OCH$_2$C<u>H</u>$_2$CH$_2$S), 2.61 (2H, t, *J* 7.2 Hz, OCH$_2$CH$_2$C<u>H</u>$_2$S), 2.82 (2H, t, *J* 6.7 Hz, SC<u>H</u>$_2$CH$_2$NH$_2$), 3.19 (2H, t, *J* 6.7 Hz, SCH$_2$C<u>H</u>$_2$NH$_2$), 3.61 (2H, t, *J* 6.1 Hz, OC<u>H</u>$_2$CH$_2$CH$_2$S), 3.65 - 3.85 (4H, m, H-3, H-4, H-5 & H-6), 3.93 (1H, dd, $J_{5,6'}$ 1.6 Hz, $J_{6,6'}$ 12.1, H-6'), 4.25 (1H, t, *J* 9.8 Hz, H-2), 4.65 (2H, s, CC<u>H</u>$_2$O), 5.84 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 8.23 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 21.6 (q, C(O)<u>C</u>H$_3$), 27.2 (t, OCH$_2$CH$_2$<u>C</u>H$_2$S), 28.1 (t, S<u>C</u>H$_2$CH$_2$NH$_2$), 28.3 (t, OCH$_2$<u>C</u>H$_2$CH$_2$S), 38.2 (t, SCH$_2$<u>C</u>H$_2$NH$_2$), 55.3 (d, C-2), 60.3 (t, C-6), 62.4 (t, C<u>C</u>H$_2$O),

68.3 (t, O<u>C</u>H$_2$CH$_2$CH$_2$S), 69.2 (d, C-4), 73.4 (d, C-3), 78.9 (d, C-5), 86.3 (d, C-1), 123.7 (d, N<u>C</u>H=C), 174.0 (s, <u>C</u>(O)CH$_3$); HRMS (ESI-TOF): calcd. for $C_{16}H_{30}N_5O_6S^+$: 420.1911. Found: 420.1922 (MH^+).

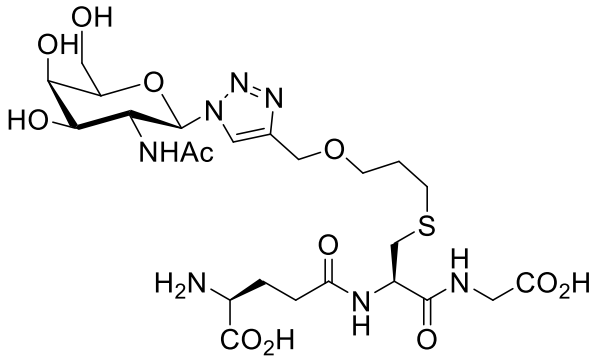
_S_-(3'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine 4.23



3-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.9** (20.5 mg, 0.06 mmol) and γ-L-glutamyl-L-cysteinylglycine **4.22** (18.4 mg, 0.06 mmol) were dissolved in sodium acetate buffer (1 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (1.5 mg, 0.006 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (water) to afford *S*-(3'-((1''-(2'''-acetamido-2'''-deoxy-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine **4.23** (38.8 mg, quant.) as a colourless oil; $[\alpha]_D^{20}$ -10 (*c*, 0.7 in water); υ_{max} 3246 cm^{-1} (OH), 1728 cm^{-1} (OC=O), 1641 (NHC=O); δ_H (400 MHz, D_2O) 1.79 (3H, s, C(O)C<u>H</u>$_3$) 1.84 (2H, quin, *J* 5.9 Hz, CH$_2$C<u>H</u>$_2$CH$_2$), 2.19 (2H, q, *J* 6.4 Hz, Q-CHC<u>H</u>$_2$), 2.50 - 2.58 (2H, m, CH$_2$C<u>H</u>$_2$S), 2.58 - 2.65 (2H, m, Q-CC<u>H</u>$_2$), 2.84 (1H, dd, *J* 13.9 Hz, *J* 8.8 Hz, C-C<u>H</u>H'), 3.01 (1H, dd, *J* 14.3 Hz, *J* 4.9 Hz, C-CH<u>H</u>'), 3.60 (1H, t, *J* 6.1 Hz, OC<u>H</u>$_2$CH$_2$), 3.65 - 3.85 (4H, m, H-3, H-4, H-5 & H-6), 3.89 - 4.01 (4H, m, G-C<u>H</u>$_2$, Q-C<u>H</u> & H-6'), 4.25 (1H, t, *J* 10.0 Hz, H-2), 4.54 (2H,

dd, *J* 8.6 Hz, *J* 5.1 Hz, C-C<u>H</u>), 4.64 (1H, s, CC<u>H</u>$_2$O), 5.84 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 8.23 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 25.7 (t, Q-CH<u>C</u>H$_2$), 28.1 (t, CH$_2$<u>C</u>H$_2$S), 28.4 (t, CH$_2$<u>C</u>H$_2$CH$_2$), 31.0 (t, Q-C<u>C</u>H$_2$), 32.7 (t, C-<u>C</u>H$_2$), 41.2 (t, G-<u>C</u>H$_2$), 53.0 (d, Q-<u>C</u>H), 53.0 (d, C-<u>C</u>H), 55.3 (d, C-2), 60.4 (t, C-6), 62.4 (t, C<u>C</u>H$_2$O), 68.3 (t, O<u>C</u>H$_2$CH$_2$), 69.2 (d, C-4), 73.4 (d, C-3), 78.8 (d, C-5), 86.3 (d, C-1), 123.7 (d, N<u>C</u>H=C), 172.5, 172.8, 172.9, 174.0, 174.4 (5 x s, 5 x <u>C</u>=O); HRMS (ESI-TOF): calcd. for $C_{24}H_{40}N_7O_{12}S^+$: 650.2450. Found: 650.2450 (MH^+).

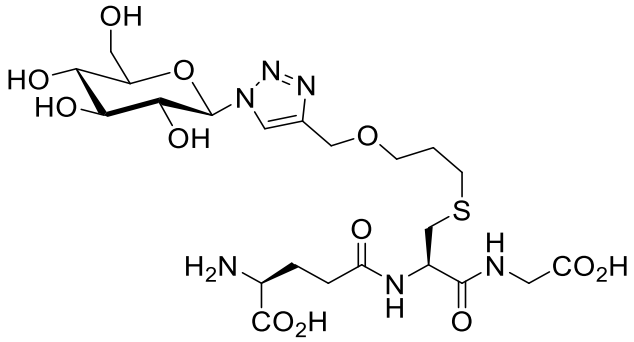
_S_-(3'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-galactopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine 4.24



3-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.10** (20.5 mg, 0.06 mmol) and γ-L-glutamyl-L-cysteinylglycine **4.22** (18.4 mg, 0.06 mmol) were dissolved in sodium acetate buffer (1 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (1.5 mg, 0.006 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (water) to afford *S*-(3'-((1''-(2'''-acetamido-2'''-deoxy-β-D-galactopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine **4.24** (27.2 mg, 70%) as a white solid; $[\alpha]_D^{20}$ -9 (*c*, 0.5 in water); υ_{max} 3362 cm^{-1}

(OH), 1637 cm^{-1} (C=O); δ_H (400 MHz, D_2O) 1.77 - 1.88 (5H, m, $CH_2C\underline{H}_2CH_2$ & $C(O)C\underline{H}_3$), 2.13 (2H, q, *J* 7.3 Hz, Q-$CHC\underline{H}_2CH_2$), 2.51 (2H, t, *J* 7.6 Hz, Q-$CHCH_2C\underline{H}_2$), 2.60 (2H, t, *J* 7.2 Hz, $CH_2C\underline{H}_2S$), 2.83 (1H, dd, *J* 14.1 Hz, *J* 9.0 Hz, C-C<u>H</u>H'), 3.03 (1H, dd, *J* 13.9 Hz, *J* 4.9 Hz, C-CH<u>H</u>'), 3.60 (2H, t, *J* 6.3 Hz, $OC\underline{H}_2CH_2$), 3.72 - 3.77 (3H, m, Q-C<u>H</u> & G-$C\underline{H}_2$), 3.79 - 3.84 (2H, m, H-6 & H-6'), 3.94 - 4.02 (2H, m, H-3 & H-5), 4.09 (1H, d, $J_{3,4}$ 2.7 Hz, H-4), 4.43 (1H, t, *J* 10.2 Hz, H-2), 4.54 (1H, dd, *J* 9.0 Hz, *J* 4.7 Hz, C-C<u>H</u>), 4.65 (2H, s, $CC\underline{H}_2O$), 5.78 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 8.27 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 21.7 (q, $C(O)\underline{C}H_3$), 26.2 (t, Q-$CH\underline{C}H_2CH_2$), 28.1 (t, $CH_2\underline{C}H_2S$), 28.4 (t, $CH_2\underline{C}H_2CH_2$), 31.4 (t, Q-$CHCH_2\underline{C}H_2$), 32.9 (t, C-$\underline{C}H_2$), 43.3 (t, G-$\underline{C}H_2$), 51.9 (d, C-2), 53.1 (d, C-$\underline{C}H$), 54.1 (d, Q-$\underline{C}H$), 60.9 (t, C-6), 62.5 (t, $C\underline{C}H_2O$), 67.7 (d, C-4), 68.3 (t, $O\underline{C}H_2CH_2$), 70.7 (d, C-3), 78.3 (d, C-5), 86.8 (d, C-1), 123.6 (d, N<u>C</u>H=C), 144.2 (s, CH=<u>C</u>CH_2), 171.9, 173.9, 174.2, 174.8, 176.1 (5 x s, 5 x <u>C</u>(O)); HRMS (ESI-TOF): calcd. for $C_{24}H_{40}N_7O_{12}S^+$: 650.2450. Found: 650.2475 (MH^+).

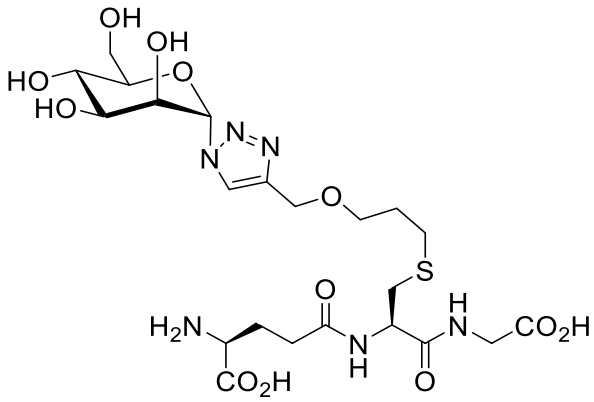
_S_-(3'-((1''-β-D-Glucopyranosyl-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine 4.25



3-((1'-β-D-Glucopyranosyl-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.11** (17.0 mg, 0.06 mmol) and γ-L-glutamyl-L-cysteinylglycine **4.22** (17.0 mg, 0.06 mmol) were dissolved in sodium acetate buffer (1 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (1.5 mg,

0.006 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (water) to afford *S*-(3'-((1''-β-D-glucopyranosyl-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine **4.25** (30 mg, 93%) as a colourless oil; $[\alpha]_D^{20}$ -23 (*c*, 1.0 in water); υ_{max} 3272 (OH), 1643, 1556 (C(O)N); δ_H (400 MHz, D_2O) 1.85 (2H, quin, *J* 7.0 Hz, $CH_2C\underline{H}_2CH_2$), 2.13 (2H, q, *J* 7.0 Hz, Q-$CHC\underline{H}_2CH_2$), 2.50 (2H, t, *J* 7.0 Hz, Q-$CHCH_2C\underline{H}_2$), 2.60 (2H, t, *J* 6.8 Hz, $CH_2C\underline{H}_2S$), 2.82 (1H, dd, *J* 14.1 Hz, *J* 9.0 Hz, C-C<u>H</u>H'), 3.03 (1H, dd, *J* 13.9 Hz, *J* 4.9 Hz, CH<u>H</u>'), 3.57 - 3.81 (9H, m, H-3, H-4, H-5, H-6, G-$C\underline{H}_2$, Q-C<u>H</u> & $OC\underline{H}_2CH_2$), 3.89 (1H, d, $J_{6,6'}$ 11.0 Hz, H-6'), 3.99 (1H, t, *J* 9.2 Hz, H-2), 4.54 (1H, dd, *J* 8.8 Hz, *J* 4.9 Hz, C-C<u>H</u>), 4.67 (2H, s, $CC\underline{H}_2O$), 5.75 (1H, d, $J_{1,2}$ 9.4 Hz, H-1), 8.26 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 26.2 (t, Q-$CH\underline{C}H_2CH_2$), 28.0 (t, $CH_2\underline{C}H_2S$), 28.4 (t, $CH_2\underline{C}H_2CH_2$), 31.4 (t, Q-$CHCH_2\underline{C}H_2$), 32.9 (t, C-$\underline{C}H_2$), 43.4 (t, G-$\underline{C}H_2$), 53.1 (d, C-<u>C</u>H), 54.1 (d, Q-<u>C</u>H), 60.4 (t, C-6), 62.6 (t, $C\underline{C}H_2O$), 68.5 (t, $O\underline{C}H_2CH_2$), 68.9 (d, C-4), 72.2 (d, C-2), 75.9 (d, C-3), 78.8 (d, C-5), 87.4 (d, C-1), 124.3 (d, N<u>C</u>H=C), 144.3 (s, NCH=<u>C</u>), 171.9, 174.9, 181.4 (3 x s, 3 x <u>C</u>(O)); HRMS (ESI-TOF): calcd. for $C_{22}H_{37}N_6O_{12}S^+$: 609.2185. Found: 609.2204 (MH^+).

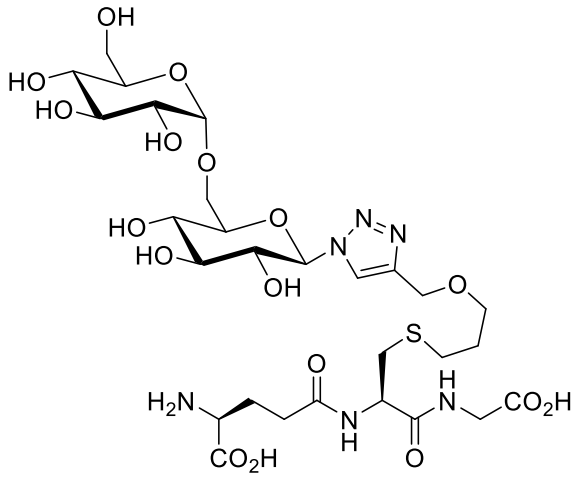
S-(3'-((1''-α-D-Mannopyranosyl-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine 4.26



3-((1'-β-D-Mannopyranosyl-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.12** (18.1 mg, 0.06 mmol) and γ-L-glutamyl-L-cysteinylglycine **4.22** (18.1 mg, 0.06 mmol) were dissolved in sodium acetate buffer (1 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (1.5 mg, 0.006 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (water) to afford *S*-(3'-((1''-α-D-mannopyranosyl-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-propyl)-γ-L-glutamyl-L-cysteinylglycine **4.26** (25.2 mg, 74%) as a colourless oil; $[\alpha]_D^{20}$ -3 (*c*, 1.0 in water); υ_{max} 3217 cm^{-1} (OH), 1634 (C=O); δ_H (400 MHz, D_2O) 1.84 (2H, quin, *J* 6.7 Hz, $CH_2C\underline{H}_2CH_2$), 2.13 (3H, q, *J* 7.3 Hz, Q-$CHC\underline{H}_2CH_2$), 2.51 (3H, t, *J* 7.8 Hz, Q-$CHCH_2C\underline{H}_2$), 2.60 (2H, t, *J* 7.2 Hz, $CH_2C\underline{H}_2S$), 2.82 (1H, dd, *J* 13.9 Hz, *J* 9.2 Hz, C-C<u>H</u>H'), 3.03 (1H, dd, *J* 14.3 Hz, *J* 4.9 Hz, C-CH<u>H</u>'), 3.25 - 3.34 (1H, m, H-5), 3.65 (2H, t, *J* 5.9 Hz, $OC\underline{H}_2CH_2$), 3.72 - 3.85 (6H, m., H-4, H-6, H-6', Q-C<u>H</u> & G-$C\underline{H}_2$), 4.13 (1H, dd, $J_{2,3}$ 3.5 Hz, $J_{3,4}$ 9.0 Hz, H-3), 4.53 (1H, dd, *J* 8.8 Hz, *J* 4.9 Hz, C-C<u>H</u>), 4.66 (2H, s, $CC\underline{H}_2O$), 4.78 - 4.82 (1H, m, H-2), 6.12 (1H, s, H-1), 8.20 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 26.1 (t, Q-$CH\underline{C}H_2CH_2$), 28.0 (t, $CH_2\underline{C}H_2S$), 28.3 (t, $CH_2\underline{C}H_2CH_2$), 31.4 (t, Q-$CHCH_2\underline{C}H_2$), 32.8 (t, C-$\underline{C}H_2$), 43.4 (t, G-$\underline{C}H_2$), 53.1 (d, C-2), 54.1 (d, Q-<u>C</u>H), 60.4 (t, C-6), 62.5 (t, $C\underline{C}H_2O$), 66.5 (d, C-4), 68.2 (d, C-2), 68.4

(t, O<u>C</u>H$_2$CH$_2$), 70.4 (d, C-3), 76.1 (d, C-5), 86.7 (d, C-1), 124.7 (d, N<u>C</u>H=C), 144.4 (s, NCH=<u>C</u>), 171.9, 173.9, 174.8 (3 x s, 3 x <u>C</u>(O)); HRMS (ESI-TOF): calcd. for $C_{22}H_{37}N_6O_{12}S^+$: 609.2185. Found: 609.2211 (MH^+).

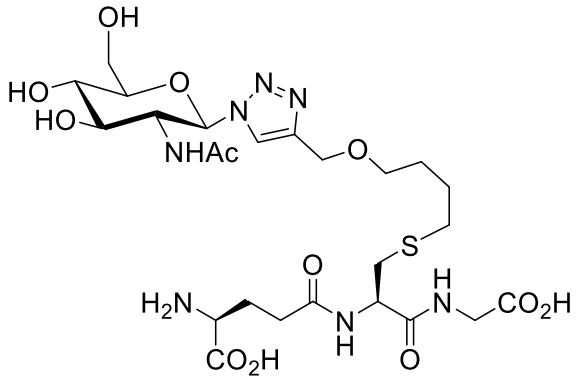
_S_-(5'-((1''-(α-D-Glucopyranosyl-(1→4)-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-pentyl)-γ-L-glutamyl-L-cysteinylglycine 4.27



3-((1'-(α-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.14** (16.0 mg, 0.035 mmol) and γ-L-glutamyl-L-cysteinylglycine **4.22** (10.6 mg, 0.035 mmol) were dissolved in sodium acetate buffer (0.5 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (0.9 mg, 0.0035 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by HPLC (t_R = 20.5 min, column: Luna 5u C18 (100 Å) column (Phenomenex); gradient: 100% water for 10 min, followed by an increase to 40% MeOH over 5 min, then 40% MeOH for 10 min; column oven: 40 °C; flow rate: 2 mL/min; detection: UV 210 nm) to afford *S*-(5'-((1''-(α-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-pentyl)-γ-L-glutamyl-L-cysteinylglycine **4.27** (20.1 mg, 76%) as a white solid; $[\alpha]_D$ +18 (*c*, 0.7 in water); υ_{max} 3306 cm^{-1} (OH), 1727 cm^{-1} (HOC=O), 1658 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.82 (2H, quin, *J* 6.7 Hz,

$OCH_2C\underline{H}_2CH_2S$), 2.12 - 2.28 (2H, m, *J* 7.0 Hz, Q-$CHC\underline{H}_2CH_2$), 2.51 - 2.63 (4H, m, $CH_2C\underline{H}_2S$ & $QCHCH_2C\underline{H}_2$), 2.81 (1H, dd, *J* 14.1 Hz, *J* 8.6 Hz, C-C<u>H</u>H'), 2.98 (1H, dd, *J* 14.1, 5.5 Hz, C-CH<u>H</u>'), 3.31 - 3.39 (1H, m, H-5_b), 3.47 (1H, dd, $J_{1,2}$ 3.5 Hz, $J_{2,3}$ 9.8 Hz, H-2_b), 3.56 - 3.71 (8H, m, H-3_a, H-5_a, H-3_b, H-4_b, H-6_b, H-6'$_b$ & $OC\underline{H}_2CH_2$), 3.81 (1H, d, $J_{6,6'}$ 10.2 Hz, H-6_a), 3.85 - 3.94 (2H, m, H-4_a & H-6'$_a$), 3.96 (2H, s, G-CH_2), 4.02 (1H, t, *J* 9.2 Hz, H-2_a), 4.06 (1H, t, *J* 6.7 Hz, Q-CH), 4.51 (1H, dd, *J* 8.4 Hz, *J* 5.3 Hz, C-CH), 4.64 (2H, s, $CC\underline{H}_2O$), 4.86 (1H, d, $J_{1,2}$ 3.5 Hz, H-1_b), 5.73 (1H, d, $J_{1,2}$ 9.4 Hz, H-1_a), 8.23 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 25.4 (t, Q-$CH\underline{C}H_2CH_2$), 28.1 (t, $CH_2\underline{C}H_2S$), 28.4 (t, Q-$CH\underline{C}H_2CH_2$), 30.8 (t, Q-$CHCH_2\underline{C}H_2$), 32.7 (t, C-$\underline{C}H_2$), 41.0 (d, G-$\underline{C}H_2$), 52.1 (t, Q-<u>C</u>H), 53.0 (d, C-<u>C</u>H), 60.0 (t, C-6_b), 62.5 (t, $C\underline{C}H_2O$), 65.9 (t, C-6_a), 68.5 (t, $O\underline{C}H_2CH_2$), 69.0 (d, C-5_a), 69.2 (d, C-5_b), 71.3 (d, C-2_b), 71.6 (d, C-4_b), 72.0 (d, C-2_a), 72.9 (d, C-3_b), 76.1 (d, C-3a), 77.5 (d, C-4_a), 87.3 (d, C-1_a), 98.0 (d, C-1_b), 124.5 (d, N<u>C</u>H=C), 144.2 (s, NCH=<u>C</u>CH_2), 171.4, 172.7, 172.8, 174.2 (4 x s, 4 x C(O)); HRMS (ESI-TOF): calcd. for $C_{28}H_{47}N_6O_{17}S^+$: 771.2713. Found: 771.2729 (MH^+).

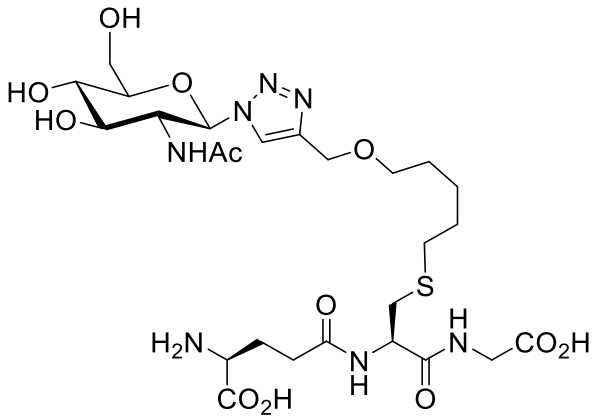
_S_-(4'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-galactopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-butyl)-γ-L-glutamyl-L-cysteinylglycine 4.28



4-((1'-(2"-acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-butene **4.26** (21.4 mg, 0.06 mmol) and γ-L-glutamyl-L-cysteinylglycine (18.4 mg, 0.06 mmol)

were dissolved in sodium acetate buffer (1 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (1.5 mg, 0.006 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (water) to afford *S*-(4'-((1''-(2'''-acetamido-2'''-deoxy-β-D-galactopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-butyl)-γ-L-glutamyl-L-cysteinylglycine **4.28** (24.9 mg, 63%) as a white solid; $[\alpha]_D^{20}$ -16 (*c*, 1.1 in water); υ_{max} 3246 cm^{-1} (OH), 1637 cm^{-1} (NHC=O); δ_H (400 MHz, D_2O) 1.54 - 1.69 (4H, m, $CH_2C\underline{H}_2CH_2CH_2$ & $CH_2CH_2C\underline{H}_2CH_2$), 1.80 (3H, s, $C(O)C\underline{H}_3$), 2.14 (2H, q, *J* 7.0 Hz, Q-$CHC\underline{H}_2CH_2$), 2.47 - 2.59 (4H, m, Q-$CHCH_2C\underline{H}_2$ & CH_2CH_2S), 2.83 (1H, dd, *J* 13.9 Hz, *J* 9.2 Hz, C-C<u>H</u>H'), 3.04 (1H, dd, *J* 13.9 Hz, *J* 4.9 Hz, C-CH<u>H</u>'), 3.49 - 3.57 (2H, m, $OC\underline{H}_2CH_2$), 3.65 - 3.85 (7H, m, H-3, H-4, H-5, H-6, Q-C<u>H</u> & G-$C\underline{H}_2$), 3.93 (1H, d, $J_{6,6'}$ 12.1 Hz, H-6'), 4.25 (1H, t, *J* 10.0 Hz, H-2), 4.54 (1H, dd, *J* 9.0 Hz, *J* 4.7 Hz, C-C<u>H</u>), 4.64 (2H, s, $CC\underline{H}_2O$), 5.84 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 8.22 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 21.6 (q, $C(O)\underline{C}H_3$), 25.2 (t, $CH_2\underline{C}H_2CH_2O$), 26.2 (t, Q-$CH\underline{C}H_2CH_2$), 27.5 (t, $OCH_2\underline{C}H_2CH_2$), 31.1 (t, $CH_2\underline{C}H_2S$), 31.4 (t, Q-$CHCH_2\underline{C}H_2$), 32.7 (t, C-$\underline{C}H_2$), 43.3 (t, G-$\underline{C}H_2$), 53.1 (d, C-<u>C</u>H), 54.1 (d, Q-<u>C</u>H), 55.3 (d, C-2), 60.4 (t, C-6), 62.4 (t, $C\underline{C}H_2O$), 69.2 (d, C-4), 69.5 (t, $O\underline{C}H_2CH_2$), 73.5 (d, C-3), 78.9 (d, C-5), 86.3 (d, C-1), 123.7 (d, N<u>C</u>H=C), 144.2 (s, CH=<u>C</u>CH_2), 171.9, 173.9, 174.0, 174.9, 176.1 (5 x s, 5 x <u>C</u>(O)); HRMS (ESI-TOF): calcd. for $C_{25}H_{42}N_7O_{12}S^+$: 664.2607. Found: 664.2623 (MH^+).

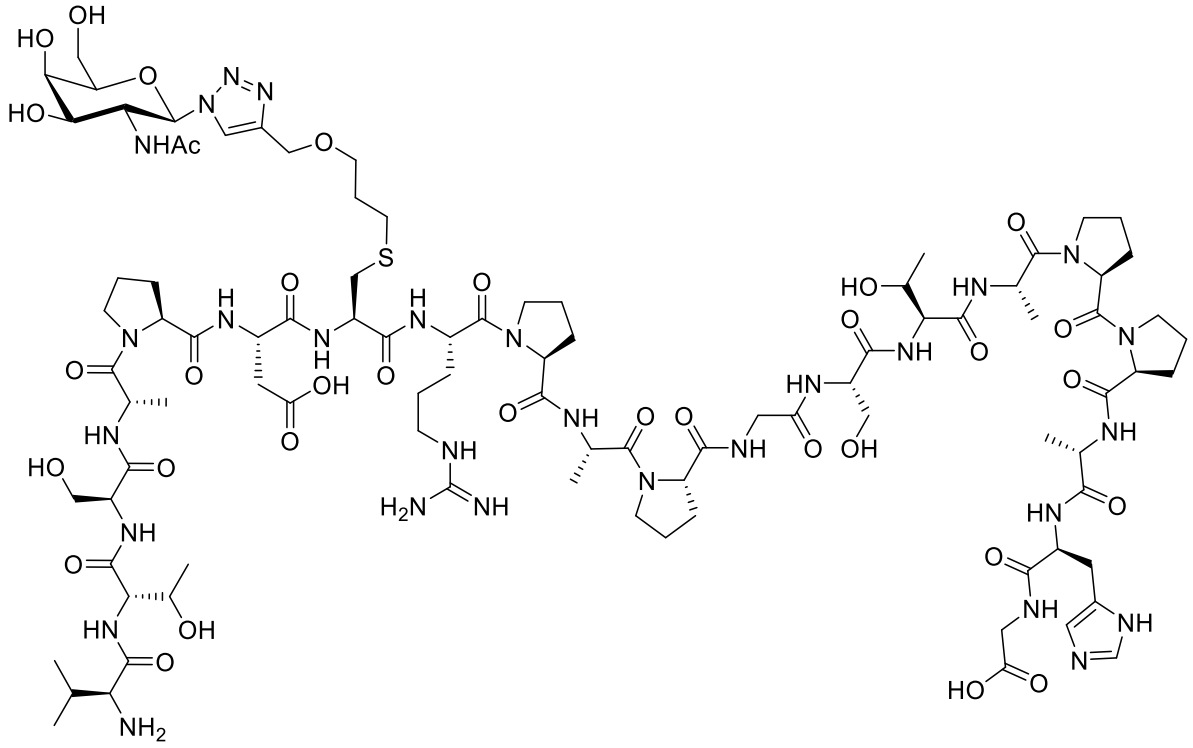
_S_-(5'-((1''-(2'''-Acetamido-2'''-deoxy-β-D-galactopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-pentyl)-γ-L-glutamyl-L-cysteinylglycine 4.29



5-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-pentene **4.27** (22.3 mg, 0.06 mmol) and γ-L-glutamyl-L-cysteinylglycine (18.4 mg, 0.06 mmol) were dissolved in sodium acetate buffer (1 mL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (1.5 mg, 0.006 mmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h. The reaction mixture was then concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (water) to afford *S*-(5'-((1''-(2'''-acetamido-2'''-deoxy-β-D-galactopyranosyl)-[1'',2'',3'']-triazo-4''-yl)methyloxy)-1'-pentyl)-γ-L-glutamyl-L-cysteinylglycine **4.29** (30.4 g, 75%) as a white solid; $[\alpha]_D^{20}$ -15 (*c*, 1.0 in water); υ_{max} 3245 cm^{-1} (OH), 1637 cm^{-1} (C=O); δ_H (400 MHz, D_2O) 1.38 (2H, quin, *J* 8.0 Hz, $OCH_2CH_2C\underline{H}_2$), 1.51 - 1.63 (4H, m, $OCH_2C\underline{H}_2CH_2$ & $CH_2C\underline{H}_2CH_2S$), 1.81 (3H, s, $NHC(O)C\underline{H}_3$), 2.15 (2H, q, *J* 7.2 Hz, Q-$CHC\underline{H}_2CH_2$), 2.49 - 2.61 (4H, m, Q-$CHCH_2C\underline{H}_2$ & $CH_2C\underline{H}_2S$), 2.85 (1H, dd, *J* 14.1 Hz, *J* 9.0 Hz, C-C<u>H</u>H'), 3.06 (1H, dd, *J* 14.1 Hz, *J* 4.7 Hz, C-CH<u>H</u>'), 3.53 (2H, t, *J* 6.7 Hz, $OC\underline{H}_2CH_2$), 3.66 - 3.86 (7H, m, H-3, H-4, H-5, H-6, G-$C\underline{H}_2$ & Q-C<u>H</u>), 3.94 (1H, d, *J* 12.1 Hz, H-6'), 4.27 (1H, t, *J* 10.0 Hz, H-2), 4.56 (1H, dd, *J* 8.6 Hz, *J* 4.7 Hz, C-C<u>H</u>), 4.65 (2H, s, $CC\underline{H}_2O$), 5.85 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 8.23 (1H, s, NC<u>H</u>=C); δ_C (100.5 MHz, D_2O) 21.6 (q, $NHC(O)\underline{C}H_3$), 24.3 (t, $OCH_2CH_2\underline{C}H_2$), 26.2 (t, Q-$CH\underline{C}H_2CH_2$), 28.0, 28.2 (2 x t,

$OCH_2\underline{C}H_2CH_2$ & $CH_2\underline{C}H_2CH_2S$), 31.4, 31.4 (2 x t, $CH_2\underline{C}H_2S$ & Q-$CHCH_2\underline{C}H_2$), 32.8 (t, C-$\underline{C}H_2$), 43.3 (t, G-$\underline{C}H_2$), 53.1 (d, C-C$\underline{H}$), 54.1 (d, Q-$\underline{C}$H), 55.3 (d, C-2), 60.4 (t, C-6), 62.3 (t, $C\underline{C}H_2O$), 69.3 (d, C-4), 69.9 (t, $O\underline{C}H_2CH_2$), 73.5 (d, C-3), 78.9 (d, C-5), 86.3 (d, C-1), 123.7 (d, N$\underline{C}$H=C), 144.3 (s, CH=$\underline{C}CH_2$), 172.0, 173.9, 174.0, 174.9, 176.1 (5 x s, 5 x $\underline{C}$(O)); HRMS (ESI-TOF): calcd. for $C_{26}H_{44}N_7O_{12}S^+$: 678.2763. Found: 678.2782 (MH^+).

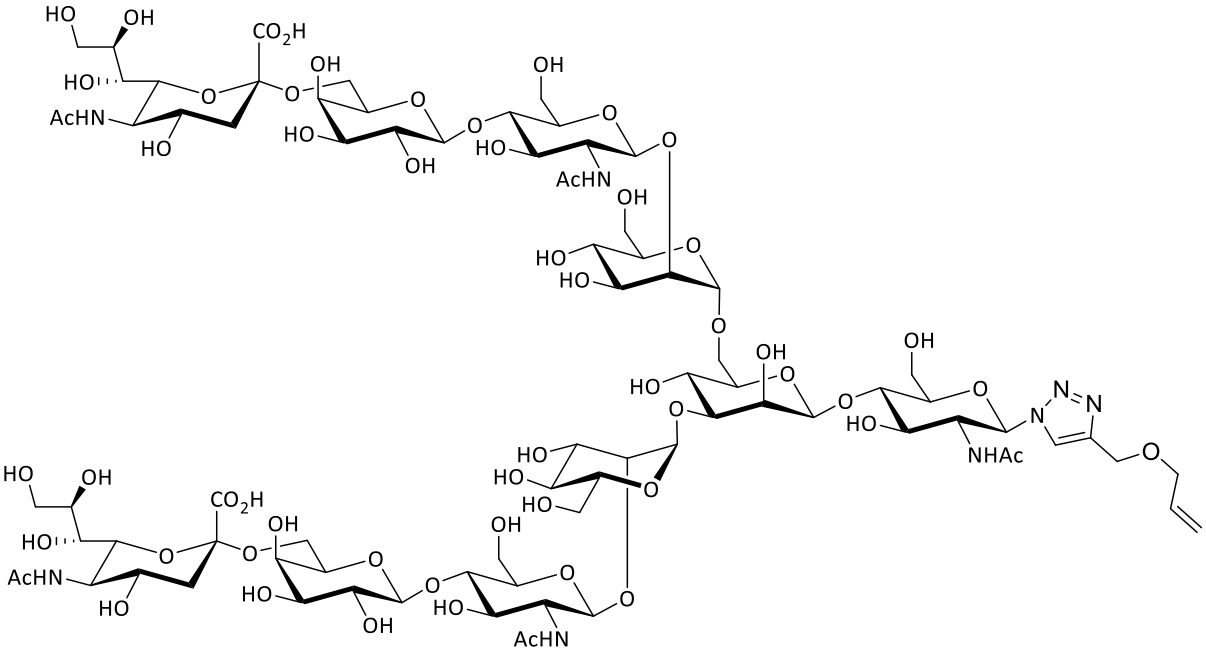
MUC1-GalNAc Neoglycopeptide 4.31



3-((1'-(2"-Acetamido-2"-deoxy-β-D-galactopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.10** (1.7 mg, 5.0 μmol) and modified MUC1 peptide **4.30** (0.9 mg, 0.5 μmol) were dissolved in sodium acetate buffer (50 μL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-phenylacetophenone (0.1 mg, 0.4 μmol) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h, and then analysed by HPLC (column: Prodigy 5u ODS3 (100 Å) column (Phenomenex); gradient: 5% MeCN in water for 5 min, followed by an increase to 25% MeCN over 20 min; column oven: 40 °C; flow rate:

0.5 mL/min; detection: UV 210 nm) to reveal the presence of MUC1-GalNAc neoglycopeptide **4.31** (80%); t_R = 19.1 min; HRMS (ESI-TOF): calcd. for $C_{93}H_{148}N_{29}O_{33}S^+$: 2231.0510. Found: 2231.0557 (MH^+).

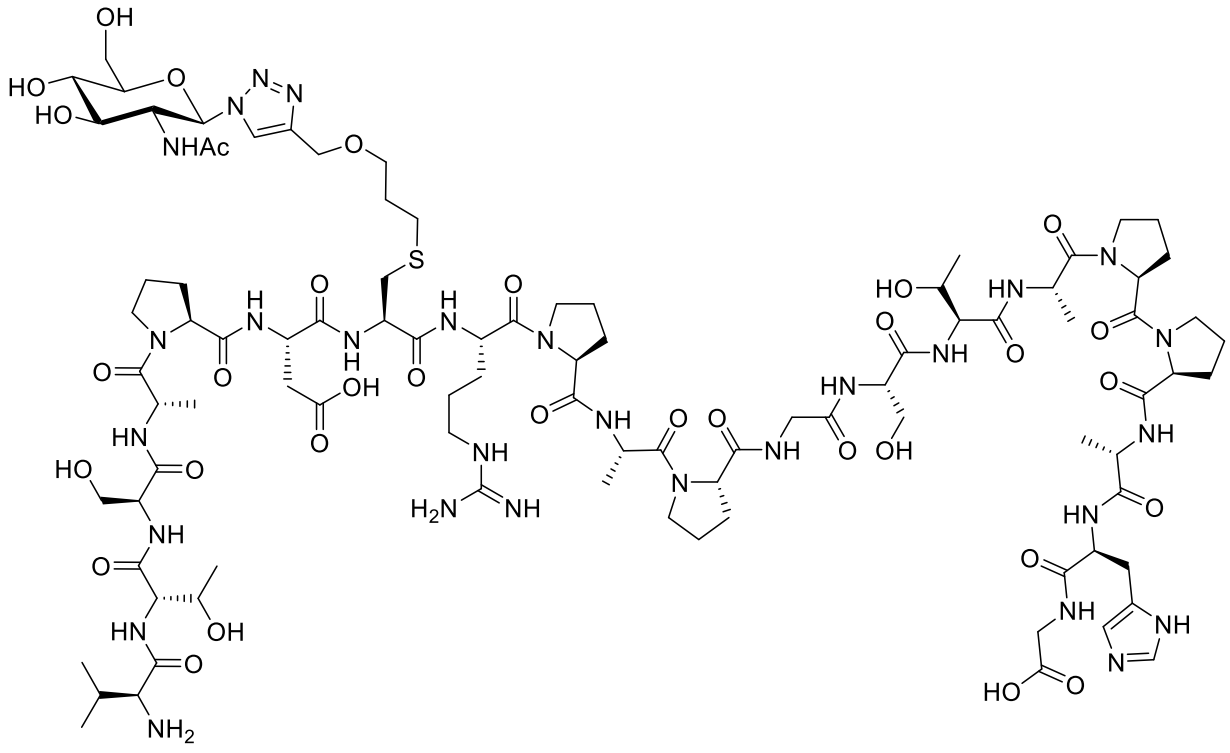
Sialoglycan-linker conjugate 4.33



Sialoglycan **4.32** (5.0 mg, 2.5 µmol) and triethylamine (24.4 µL, 175 µmol) were dissolved in D_2O/MeCN (4:1, 90 µL) and the solution cooled to 0 °C. 2-Azido-1,3-dimethylimidazolinium hexafluorophosphate **4.7** (21.4 mg, 75 µmol) was added and the mixture was stirred for 6 h. The reaction mixture was then acidified to pH 2 by addition of aqueous HCl over 10 min (1 M, 150 µL). After 20 min the reaction mixture was neutralised by the addition of $NaHCO_3$ (sat. aqueous soln., 150 µL). The mixture was lyophilised and the residue was dissolved in *tert*-butanol/water (2:1, 150 µL) and 3-allyloxy-1-propyne **4.3** (5.6 µL, 50 µmol), $CuSO_4.5H_2O$ (1.2 mg, 5 µmol), and sodium ascorbate (2.0 mg, 10 µmol) were added. The reaction mixture was then heated to 50 °C and stirred for 40 h. The reaction was then filtered and the mixture purified by HPLC (t_R = 14.5 min; column: XBridge® Glycan BEH Amide 3.5µm (130 Å)

column; gradient: 10% water in MeCN for 7 min, followed by an increase to 40% water over 5 min, followed by 40% water in MeCN for 5 min; column oven: 40 °C; flow rate: 1 mL/min; detection: UV 210 nm) to afford sialoglycan-linker conjugate **4.33** (5.0 mg, 94%) as a white solid; $[\alpha]_D^{20}$ -10 (*c*, 0.25 in water); δ_H (400 MHz, D_2O) 4.05 (2H, d, *J* 5.9 Hz, CH_2=CHC$\underline{H}_2$), 4.67 (2H, s, NCH=CC$\underline{H}_2$), 5.26 (1H, d, J_Z 10.2 Hz, $CH_E\underline{H}_Z$=CH), 5.32 (1H, d, J_E 17.2 Hz, C$\underline{H}_E H_Z$=CH), 5.82 - 6.01 (2H, m, CH_2=C$\underline{H}$ & H-1), 8.24 (1H, m, NC$\underline{H}$=C); HRMS (ESI-TOF): calcd. for $C_{82}H_{133}N_8O_{57}^+$: 2141.7749. Found: 2141.7770 (MH^+).

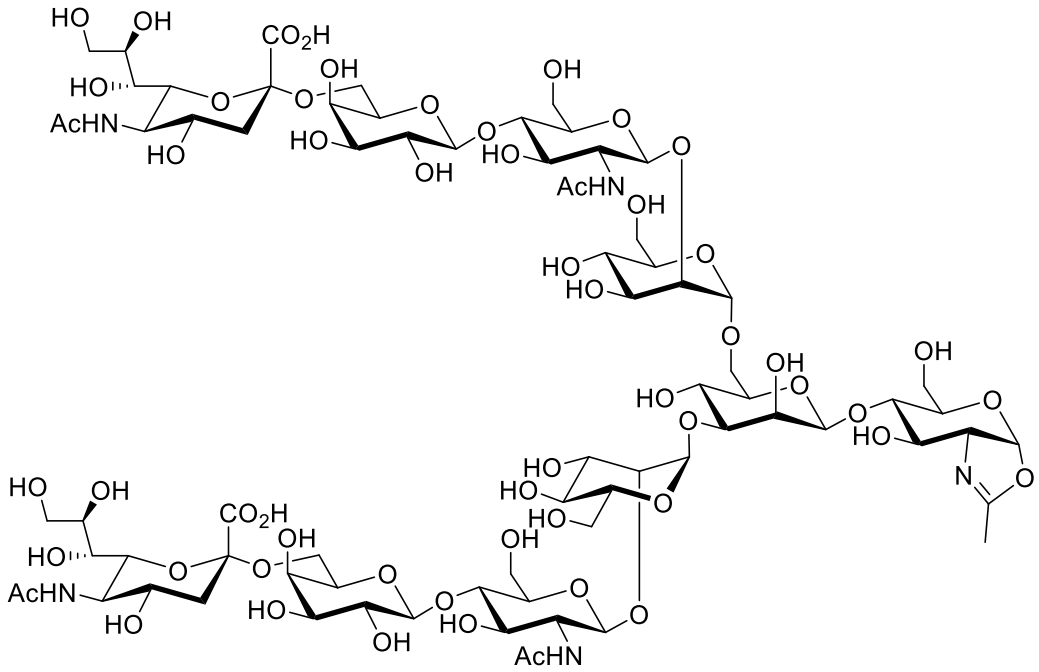
MUC1-GlcNAc Neoglycopeptide 4.40



3-((1'-(2"-Acetamido-2"-deoxy-β-D-glucopyranosyl)-[1',2',3']-triazo-4'-yl)methyloxy)-1-propene **4.9** (0.45 mg, 0.13 μmol) and modified MUC1 peptide **4.30** (0.25 mg, 0.0.13 μmol) were dissolved in sodium acetate buffer (15 μL, pH 4.0, 0.2 M). 2,2-Dimethoxy-2-

phenylacetophenone (33 µg, 0.13 µmol, added as a solution in 1 µL DMF) was added, and the reaction mixture was stirred and irradiated in a Rayonet photoreactor (wavelength: 254 nm) for 4 h, and then analysed by HPLC (column: Jupiter 5u C4 (300 Å) column (Phenomenex); gradient: 5% MeCN in water for 5 min, followed by an increase to 25% MeCN over 20 min; column oven: 40 °C; flow rate: 0.5 mL/min; detection: UV 210 nm) to reveal the presence of MUC1-GlcNAc neoglycopeptide **4.35** (77%); t_R = 16.0 min; HRMS (ESI-TOF): calcd. for $C_{93}H_{148}N_{29}O_{33}S^+$: 2231.0510. Found: 2231.0532 (MH^+).

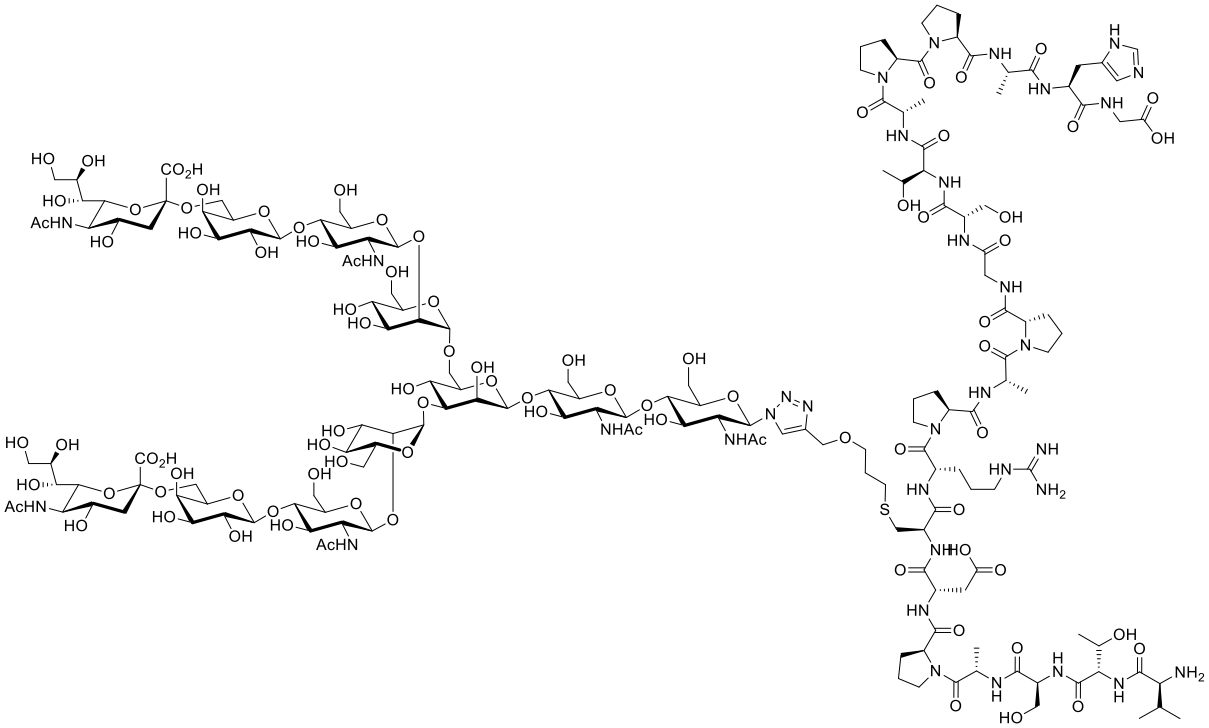
Sialoglycan oxazoline 4.41[24]



Sialoglycan **4.32** (5.0 mg, 2.5 µmol) and triethylamine (35 µL, 250 µmol) were dissolved in water (100 µL) and the solution cooled to 0 °C. 2-Chloro-1,3-dimethylimidazolinium chloride **2.1** (12.7 mg, 75 µmol) was added and the mixture was stirred for 6 h. The reaction was then purified by HPLC (t_R = 14.5 min; column: XBridge® Glycan BEH Amide 3.5µm (130 Å) column; gradient: 10% water in MeCN for 7 min, followed by an increase to 40% water over

5 min, followed by 40% water in MeCN for 5 min; column oven: 40 °C; flow rate: 1 mL/min; detection: UV 210 nm) to afford sialoglycan oxazoline **4.41** (3.1 mg, 63%) as a white solid; δ_H (400 MHz, D_2O) 2.01 (3H, s, CH_3), 2.04 (12H, s, 4 x CH_3), 4.94 (1H, s, 1 x H-1), 5.10 (1H, s, 1 x H-1), 6.07 (1H, d, $J_{1,2}$ 7.0 Hz, H-1_a); HRMS (ESI-TOF): calcd. for $C_{76}H_{124}N_5O_{56}^+$: 2002.7003. Found: 2002.7037 (MH^+).

MUC1-Sialoglycan Neoglycopeptide 4.42[25]

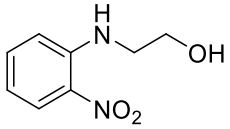


MUC1-GlcNAc Neoglycopeptide **4.40** (0.2 mg, 0.1 µmol) and sialoglycan oxazoline **4.41** (0.8 mg, 0.4 µmol) were dissolved in DMSO (4 µL). The mixture was heated to 37 °C and Endo M N175Q (10 mU, solution in sodium phosphate buffer (10 µL, pH 6.5, 0.1 M)) was added and the reaction was heated at 37 °C. Analysis by HPLC (column: Jupiter 5u C4 (300

Å) column (Phenomenex); gradient: 5% MeCN in water for 5 min, followed by an increase to 25% MeCN over 20 min; column oven: 40 °C; flow rate: 0.5 mL/min; detection: UV 210 nm) indicated complete reaction after 1 h, and the presence of MUC1-sialoglycan neoglycopeptide **4.42** (40%); t_R = 15.1 min; HRMS (ESI-TOF): calcd. for $C_{169}H_{271}N_{34}O_{89}S^+$: 4232.7740. Found: 4232.7307 (MH^+).

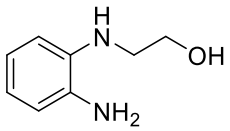
Experimental section for Chapter 5

2-Nitro-*N*-(2-hydroxyethyl)aniline 5.4[26]



1-Chloro-2-nitrobenzene (0.79 g, 5 mmol) was dissolved in DMSO (10 mL) containing K_2CO_3 (0.82 g, 6 mmol) under an atmosphere of nitrogen. Ethanolamine (0.48 mL, 15 mmol) was added and the reaction refluxed for 16 h, at which point t.l.c. (CH_2Cl_2:ethyl acetate, 2:1) indicated complete consumption of starting material (R_f 0.9) and formation of a single product (R_f 0.6). The reaction mixture was then cooled and diluted with water (50 mL) and then brine (50 mL). The resulting solution was extracted with Et_2O (3 x 50 mL) and the combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:ethyl acetate, 2:1) to afford 2-nitro-*N*-(2-hydroxyethyl)aniline **5.4** (0.53 g, 58%) as an orange solid; m.p. 68-70 °C (EtOH) [lit. 71-73 °C (EtOH)][27]; δ_H (400 MHz, $CDCl_3$)[26] 3.49 (2H, q, *J* 5.5 Hz, NHC$\underline{H}_2$CH$_2$OH), 3.93 (2H, t, *J* 5.5 Hz, NHCH$_2$C$\underline{H}_2$OH), 6.64 (1H, t, *J* 7.8 Hz, Ar-H), 6.88 (1H, d, *J* 8.6 Hz, Ar-H), 7.42 (1H, t, *J* 8.2 Hz, Ar-H), 8.13 (1H, d, *J* 8.6 Hz, Ar-H); δ_C (100.5 MHz, $CDCl_3$)[26] 45.0 (t, NH$\underline{C}$H$_2$CH$_2$OH), 60.8 (t, NHCH$_2\underline{C}$H$_2$OH), 113.8 (d, Ar-C), 115.6 (d, Ar-C), 126.9 (d, Ar-C), 132.1(s, Ar-C), 136.3 (d, Ar-C), 145.5 (s, Ar-C); HRMS (ESI-TOF): calcd. for $C_8H_{11}O_3N_2^+$: 183.0764. Found: 183.0766 (MH^+).

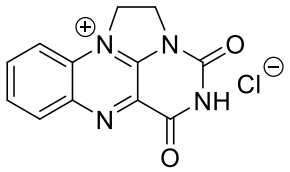
2-((2-Hydroxyethyl)amino)aniline 5.5[26]



2-Nitro-*N*-(2-hydroxyethyl)aniline **5.4** (1.00 g, 5.49 mmol) and 10% Pd on carbon (0.92 g) were dissolved in MeOH (20 mL) under an atmosphere of nitrogen. The mixture was cooled

to 0 °C and ammonium formate (1.73 g, 27.4 mmol) was added. The reaction was then warmed to room temperature and stirred for 1 h. At this point t.l.c. (CH_2Cl_2:ethyl acetate, 2:1) showed complete consumption of starting material (R_f 0.6) and formation of a single product (R_f 0.1). The reaction mixture was then filtered through Celite®, washing with MeOH. The filtrate was concentrated *in vacuo*. Brine (25 mL) was added and the mixture was extracted with CH_2Cl_2 (3 x 20 mL). The organic extracts were combined, dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford 2-((2-hydroxyethyl)amino)aniline **5.5** (0.55 g, 66%) as a white solid; δ_H (400 MHz, DMSO-d_6)[26] 3.07 (2H, q, *J* 5.9 Hz, NHC$\underline{H}_2$CH$_2$OH), 3.59 (2H, q, *J* 5.6 Hz, NHCH$_2$C$\underline{H}_2$OH), 4.30 (1H, m, O$\underline{H}$), 4.40 (2H, s, N$\underline{H}_2$), 4.66 (1H, t, *J* 5.6 Hz, N$\underline{H}$), 6.38-6.55 (4H, m, 4 x Ar-H).

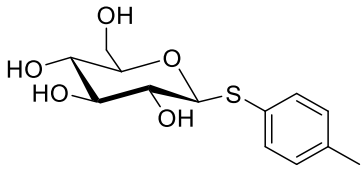
1,10-Ethyleneisoalloxazinium chloride 5.7[26]



2-((2-Hydroxyethyl)amino)aniline **5.5** (0.50 g, 3.28 mmol) was dissolved in glacial acetic acid (10 mL). The resulting solution was heated to 50 °C in the dark and alloxan monohydrate **5.6** (0.63 g, 3.94 mmol) was added, followed by boric acid (0.243 g, 3.94 mmol). The reaction was stirred for 16 h and then allowed to cool. The resulting precipitate was filtered off, washed with CH_2Cl_2 and dried under vacuum. The solid was then placed in the dark under an atmosphere of nitrogen. Thionyl chloride (7.5 mL) was added gradually, with constant stirring. The solution was then heated to 50 °C and stirred for 16 h. After this time the reaction was cooled, and the precipitate filtered off and washed with CH_2Cl_2 (12.5 mL). The crude product was dissolved in the minimum amount of 98% formic acid and then precipitated by the addition of Et_2O to afford 1,10-ethyleneisoalloxazinium chloride **5.7** (0.29 g, 32%) as a brown solid;

m.p. 238-240 °C (decomp.) [lit. 216-219 °C][26]; δ_H (400 MHz, D_2O)[26] 4.77 (2H, t, *J* 9.6 Hz, C<u>H</u>$_2$), 5.43 (2H, t, *J* 9.4 Hz, C<u>H</u>$_2$), 8.03-8.13 (2H, m, 2 x Ar-H), 8.33 (1H, t, *J* 7.8 Hz, Ar-H), 8.48 (1H, d, *J* 8.6 Hz, Ar-H); δ_C (100.5 Hz, D_2O) 45.0 (t, <u>C</u>H_2), 51.2 (t, <u>C</u>H_2), 117.1 (d, Ar-C), 129.9 (s, Ar-C), 132.1 (d, Ar-C), 132.7 (d, Ar-C), 140.6 (d, Ar-C), 140.8 (s, Ar-C), 143.5 (s, Ar-C), 159.0 (s, Ar-C); HRMS (ESI-TOF): calcd. for $C_{12}H_9O_2N_4^+$: 241.0726. Found: 241.0720 (M^+).

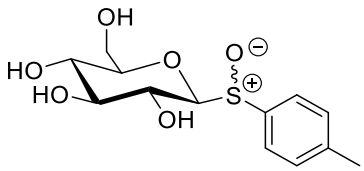
_p_-Tolyl-1-thio-β-D-glucopyranoside 5.9



NaOMe (0.54 g, 10 mmol) was added to a stirred solution of *p*-tolyl-2,3,4,6-*O*-acetyl-1-thio-β-D-glucopyranoside **5.8** (6.82 g, 15 mmol) in MeOH (70 mL). After 1, h t.l.c. (ethyl acetate:MeOH 9:1) indicated complete consumption of starting material (R_f 0.9) and formation of a single product (R_f 0.3). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until the pH was neutral. The mixture was then filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (ethyl acetate:MeOH 9:1) to afford *p*-tolyl-1-thio-β-D-glucopyranoside **5.9** (3.07 g, 72%) as a white solid; m.p. 147-148 °C (EtOH) [lit. 147-148 °C (EtOH)][28]; $[\alpha]_D^{20}$ -45 (*c*, 1.0 in MeOH) [lit. $[\alpha]_D^{20}$ -48 (*c*, 1.0 in MeOH)][29] ; δ_H (400 MHz, DMSO-d_6)[29] 2.26 (3H, s, C<u>H</u>$_3$), 2.95-3.09 (2H, m, H-2 & H-4), 3.13-3.22 (2H, m, H-3 & H-5), 3.42 (1H, dt, $J_{6,6'}$ 11.7 Hz, *J* 5.6 Hz, H-6), 3.66 (1H, dd, $J_{6,6'}$ 11.7 Hz, $J_{6',OH-6}$ 5.5 Hz), 4.50 (2H, m, H-1 & OH-6), 4.92 (1H, d, $J_{4,OH-4}$ 5.1 Hz, OH-4), 5.03 (1H, d, $J_{3,OH-3}$ 4.7 Hz, OH-3), 5.19 (1H, d, $J_{2,OH-2}$ 5.9 Hz, OH-2), 7.11 (2H, d, *J* 8.2 Hz, 2 x Ar-H), 7.36 (2H, d, *J* 7.8 Hz, 2 x Ar-H); δ_C (100.5 MHz, DMSO-d_6) 21.0 (q, <u>C</u>H_3), 61.5 (t, C-6), 70.3 (d, C-4), 72.8 (d, C-2), 78.6 (d, C-3), 81.4 (d, C-5), 87.9 (d,

C-1), 129.9 (d, Ar-C), 131.0 (d, Ar-C), 131.3 (s, Ar-C), 136.5 (s, Ar-C); HRMS (ESI-TOF): calcd. for $C_{13}H_{18}O_5SNa^+$: 309.0767. Found: 309.0768 (MNa^+).

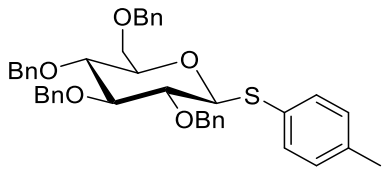
_p_-Tolyl-1-sulfinyl-β-D-glucopyranoside 5.10



_p_-Tolyl-1-thio-β-D-glucopyranoside **5.9** (0.858 g, 3 mmol) was dissolved in MeOH (8.1 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0417 g, 0.15 mmol) was added followed by H_2O_2 (0.21 mL, 50% w/w in water, 3.6 mmol), and the reaction was stirred for 6 h. The reaction mixture was concentrated *in vacuo* and the residue was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford _p_-tolyl-1-sulfinyl-β-D-glucopyranosid**e 5.10** (0.669 g, 74%) as a pale orange solid (1:0.8 mixture of diastereoisomers); m.p. 70-71 °C (MeOH); υ_{max} (neat) 1011 cm^{-1} (S=O) δ_H (400 MHz, DMSO-d_6) 2.35 (5.4H, s, C$\underline{H}_{3\,a}$ & C$\underline{H}_{3\,b}$), 2.86 (0.8H, m, H-3_a), 2.97 (1H, m, H-5_b), 3.05-3.12 (1.8H, m, H-2_a & H-4_b), 3.13-3.29 (2.6H, m, H-4_a, H-5_a & H-3_b), 3.29-3.45 (2.8H, m, H-6_a, H-6_b & H-$6'_b$), 3.49 (1H, m, H-2_b), 3.64 (0.8H, m, H-$6'_a$), 3.90 (1H, d, $J_{1,2}$ 9.8 Hz, H-1_b), 4.09 (1H, t, J 5.5 Hz, OH-6_b), 4.3 (0.8H, d, $J_{1,2}$ 9.4 Hz, H-1_a), 4.47 (0.8H, t, J 5.7 Hz, OH-6_a), 4.95 (1.8H, d, J 5.1 Hz, OH-3_a & OH-4_b), 5.05 (0.8H, d, J 5.1 Hz, OH-4_a), 5.10 (0.8H, d, J 5.5 Hz, OH-2_a), 5.15 (1H, d, J 5.5 Hz, OH-3_b), 5.68 (1H, d, J 5.9 Hz, OH-2_b), 7.32 (3.6H, d, J 7.8 Hz, 2 x Ar-H_a & 2 x Ar-H_b), 7.52 (2H, d, J 7.8 Hz, 2 x Ar-H_b), 7.55 (1.6H, d, J 7.8 Hz, 2 x Ar-H_a); δ_C (100.5 MHz, DMSO-d_6) 21.4 (q, $\underline{C}H_{3\,a}$ & $\underline{C}H_{3\,b}$), 61.2 (t, C-6_b), 61.3 (t, C-6_b), 69.0 (d, C-2_b), 69.8 (d, C-2_a), 69.9 (d, C-3_a), 69.9 (d, C-4_b), 78.1 (d, C-5_a), 78.3 (d, C-3_b), 81.7 (d, C-5_b), 82.0 (d, C-4_a), 93.6 (d, C-1_b), 94.6 (d, C-1_a), 126.0 (d, 2 x Ar-C_a), 126.4 (d, 2 x Ar-C_b), 129.5 (d, 2 x Ar-C_a), 129.7

(d, 2 x Ar-C_b), 137.4 (s, Ar-C_a), 138.0 (s, Ar-C_b), 140.9 (s, Ar-C_b), 141.2 (s, Ar-C_a); HRMS (ESI-TOF): calcd. for $C_{13}H_{19}O_6S^+$: 303.0897. Found: 303.0911 (MH^+).

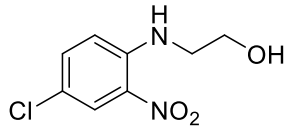
_p_-Tolyl-2,3,4,6-_O_-benzyl-1-thio-β-D-glucopyranoside 5.11



p-Tolyl-1-thio-β-D-glucopyranoside **5.9** (2.86 g, 10 mmol) was dissolved in anhydrous DMF (35 mL) under an atmosphere of nitrogen. The solution was cooled to 0 °C and NaH (60% dispersion in mineral oil, 3.2 g, 80 mmol) was slowly added to the reaction mixture, followed by portionwise addition of benzyl bromide (7.2 mL, 60 mmol). The reaction mixture was then warmed to room temperature and stirred for 16 h, at which point t.l.c. (petrol:ethyl acetate 8:1) indicated complete consumption of starting material (R_f 0), and the formation of a single product (R_f 0.4). MeOH (30 mL) was added slowly to quench the reaction. The reaction mixture was concentrated *in vacuo*, and the residue was dissolved in Et_2O (100 mL) and washed with water (30 mL). The aqueous layer was extracted with Et_2O (2 x 30 mL) and the combined organic extracts were washed with brine (40 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (petrol:ethyl acetate 8:1) to afford *p*-tolyl-2,3,4,6-*O*-benzyl-1-thio-β-D-glucopyranoside **5.11** (5.13 g, 79%) as a colourless solid; m.p. 79-80 °C (EtOH) [lit. 78-79 °C][30]; $[\alpha]_D^{20}$ -2.5 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ -1.8 (*c*, 1.0 in $CHCl_3$)][30]; δ_H (400MHz, $CDCl_3$) 2.32 (3H, s, CH_3), 3.50 (2H, m, H-2 & H-3), 3.61-3.83 (5H, m, H-1, H-4, H-5, H-6 & H-6'), 4.52-4.65 (4H, m, H-1 & 3 x C<u>H</u>H'Ph), 4.75 (1H, d, *J* 10.2 Hz, C<u>H</u>H'Ph), 4.81-4.95 (4H, m, 4 x CH<u>H</u>'Ph), 7.05 (2H, d, *J* 7.8 Hz, Ar-H), 7.19-7.45 (20H, m, 20 x Ar-H), 7.51 (2H, d, *J* 8.2 Hz, Ar-H); δ_C (100.5 MHz, $CDCl_3$) 21.1 (q, <u>C</u>H_3), 69.1 (t, C-6), 73.4 (t, <u>C</u>H_2Ph), 75.0 (t, <u>C</u>H_2Ph), 75.4 (t, <u>C</u>H_2Ph), 75.8 (t,

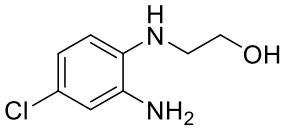
CH_2Ph), 77.8 (d, C-4), 79.1 (d, C-2), 80.8 (d, C-2), 86.8 (d, C-5), 87.8 (d, C-1), 127.5, 127.7, 127.7, 127.8, 127.8, 127.9, 128.2, 128.3, 128.4, 128.4, 129.6, 132.7, 137.7, 138.1, 138.1, 138.4, 138.4, (30 x Ar-C); HRMS (ESI-TOF): calcd. for $C_{41}H_{42}O_5SNa^+$: 669.2645. Found: 669.2658 (MNa^+).

2-Nitro-4-chloro-*N*-(2-hydroxyethyl)aniline 5.15



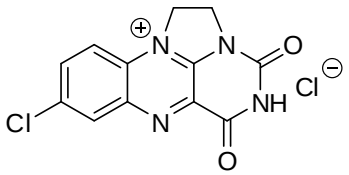
1,4-Dichloro-2-nitrobenzene **5.14** (24 g, 125 mmol) was dissolved in DMSO (250 mL) containing K_2CO_3 (20.5 g, 150 mmol) under an atmosphere of nitrogen. Ethanolamine (12 mL, 375 mmol) was added and the reaction was refluxed for 16 h, at which point t.l.c. (CH_2Cl_2:ethyl acetate, 2:1) indicated complete consumption of starting material (R_f 0.9) and formation of a single product (R_f 0.6). The reaction mixture was then cooled and diluted with water (250 mL) and then brine (250 mL). The resulting solution was extracted with Et_2O (3 x 250 mL) and the combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was crystallised from toluene to afford 2-nitro-4-chloro-*N*-(2-hydroxyethyl)aniline **5.15** (12.30 g, 45%) as an orange solid; m.p. 105-107 °C (toluene) [lit. 107.5 °C (*n*-butanol)][31]; δ_H (400 MHz, $CDCl_3$) 3.49 (2H, t, *J* 5.1 Hz, NHCH_2CH_2OH), 3.94 (2H, t, *J* 5.5 Hz, NHCH_2CH_2OH), 6.86 (1H, d, *J* 9.4 Hz, Ar-H), 7.38 (1H, d, *J* 9 Hz, Ar-H), 8.16 (1H, s, Ar-H); δ_C (100.5 MHz, $CDCl_3$) 45.0 (t, NHCH_2CH_2OH), 60.9 (t, NHCH_2CH_2OH), 115.2 (d, Ar-C), 120.4 (s, Ar-C), 126.0 (d, Ar-C), 136.3 (d, Ar-C), 144.1 (s, Ar-C), 181.1 (s, Ar-C); HRMS (ESI-TOF): calcd. for $C_8H_{10}ON_2Cl^+$: 217.0374. Found: 217.0373 (MH^+).

2-((2-Hydroxyethyl)amino)-5-chloroaniline 5.16



2-Nitro-4-chloro-*N*-(2-hydroxyethyl)aniline **5.15** (1.05g, 5 mmol) and tin (1.78 g, 15 mmol) were dissolved in water (15 mL) under an atmosphere of nitrogen. The solution was then heated to 100 °C and conc. HCl (9 mL) was added dropwise. The reaction was then cooled to 5 °C and stirred for 0.5 h. The pH was then increased to above 7 by the addition of 50% aqueous NaOH (5 mL) and the mixture was extracted with ethyl acetate (3 x 15 mL). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford 2-((2-hydroxyethyl)amino)-5-chloroaniline **5.16** (0.9196 g, 99%) as a pale pink solid; δ_H (400 MHz, DMSO-d_6) 3.05 (2H, t, *J* 5.9 Hz, NHC$\underline{H}_2$CH_2OH), 3.57 (2H, t, *J* 5.9 Hz, NHCH_2C$\underline{H}_2$OH), 6.37 (1H, m, Ar-H), 6.46 (1H, m, Ar-H), 6.54 (1H, m, Ar-H).

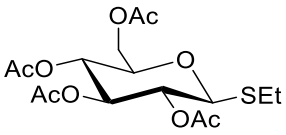
7-Chloro-1,10-ethyleneisoalloxazinium chloride 5.17



2-((2-Hydroxyethyl)amino)-5-chloroaniline **5.16** (9.89 g, 53 mmol) was dissolved in glacial acetic acid (160 mL). The resulting solution was heated to 50 °C in the dark and alloxan monohydrate **5.6** (10.2 g, 63.8 mmol) was added, followed by boric acid (3.93 g, 63.7 mmol). The reaction was stirred for 16 h and then allowed to cool. The resulting precipitate was filtered off, washed with CH_2Cl_2 and dried under vacuum. The solid was then placed in the dark under an atmosphere of nitrogen. Thionyl chloride (120 mL) was added gradually, with constant stirring. The solution was then heated to 50 °C and stirred for 16 h. After this time the reaction was cooled, and the precipitate filtered off and washed with CH_2Cl_2 (200 mL). The crude

product was dissolved in the minimum amount of 98% formic acid and then precipitated by the addition of Et_2O to afford 7-chloro-1,10-ethyleneisoalloxazinium chloride **5.17** (15.43 g, 94%) as a yellow solid; m.p. 170-200 °C (decomp.), HRMS (ESI-TOF): calcd. for $C_{12}H_8O_2N_4Cl^+$: 275.0336. Found: 275.0327 (M^+).

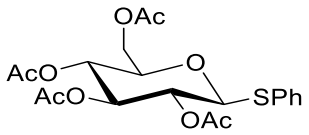
Ethyl 2,3,4,6-acetyl-1-thio-β-D-glucopyranoside 5.19[32]



1,2,3,4,6-Penta-*O*-acetyl-β-D-glucopyranose **5.18** (5.86 g, 15 mmol) was dissolved in CH_2Cl_2 (35 mL) and the resulting solution was cooled to 0 °C. Ethanethiol (1.98 mL, 27.5 mmol) was added, followed by dropwise addition of $BF_3.OEt_2$ (3.39 mL, 27.5 mmol) with constant stirring. The reaction was then warmed to room temperature, and after 2 h t.l.c. (petrol:Et_2O 1:1) indicated complete consumption of starting material (R_f 0.2) and the formation of a single product (R_f 0.3). The reaction mixture was diluted with CH_2Cl_2 (40 mL), washed with aqueous $NaHCO_3$ (sat. aqueous soln., 3 x 40 mL), followed by brine (2 x 20 mL) and the organic layer was then dried ($MgSO_4$), filtered and concentrated *in vacuo*. The resulting solid was recrystallised from the minimum amount of hot EtOH to afford ethyl 2,3,4,6-acteyl-1-thio-β-D-glucopyranoside **5.19** (3.67 g, 63%) as a colourless crystalline solid; m.p. 79-80 °C (EtOH) [lit: 81-82 °C (EtOH)][33]; $[\alpha]_D^{20}$ -28 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ -27.1 (*c*, 1.0 in $CDCl_3$)][34]; δ_H (400 MHz, $CDCl_3$)[33] 1.24 (3H, t, *J* 7.4 Hz, $CH_2C\underline{H}_3$), 1.98 (3H, s, C(O)C$\underline{H}_3$), 2.00 (3H, s, C(O)C$\underline{H}_3$), 2.03 (3H, s, C(O)C$\underline{H}_3$), 2.05 (3H, s, C(O)C$\underline{H}_3$), 2.68 (2H, m, $C\underline{H}_2CH_3$), 3.69 (1H, m, H-5), 4.11 (1H, m, H-6), 4.22 (1H, m, H-6'), 4.47 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 5.00 (1H, t, *J* 9.5 Hz, H-2), 5.05 (1H, t, 9.5 Hz, H-4), 5.20 (1H, t, *J* 9.4 Hz, H-3); δ_C (100.5 MHz, $CDCl_3$) 14.8 (q, $CH_2\underline{C}H_3$), 20.5 (q, C(O)$\underline{C}H_3$), 20.6 (q, C(O)$\underline{C}H_3$), 20.7 (q, C(O)$\underline{C}H_3$), 20.7 (q, C(O)$\underline{C}H_3$), 24.1 (t, $\underline{C}H_2CH_3$), 62.1 (t, C-6), 68.3 (d, C-4), 69.8 (d, C-2), 73.9 (d, C-3), 75.8 (d,

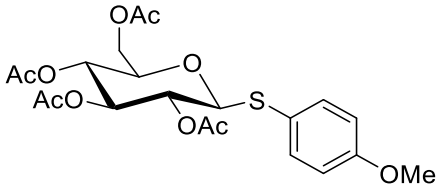
C-5), 83.5 (d, C-1), 169.3 (s, C(O)CH_3), 169.4 (s, C(O)CH_3),170.1 (s, C(O)CH_3), 170.6 (s, C(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{16}H_{24}O_9SNa^+$: 415.1033. Found: 415.1035 (MNa^+).

Phenyl 2,3,4,6-acetyl-1-thio-β-D-glucopyranoside 5.20[35]



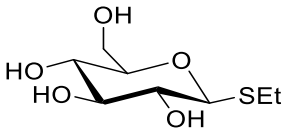
1,2,3,4,6-Penta-*O*-acetyl-β-D-glucopyranose **5.18** (5.86 g, 15 mmol) was dissolved in CH_2Cl_2 (35 mL) and the resulting solution was cooled to 0 °C. Thiophenol (2.81 mL, 27.5 mmol) was added, followed by the dropwise addition of $BF_3.OEt_2$ (3.39 mL, 27.5 mmol) with constant stirring. The reaction was then warmed to room temperature, and after 2 h t.l.c. (petrol:Et_2O 1:1) indicated complete consumption of starting material (R_f 0.5) and formation of a single product (R_f 0.6). The reaction mixture was diluted with CH_2Cl_2 (40 mL), washed with aqueous $NaHCO_3$ (sat. aqueous soln., 3 x 40 mL), followed by brine (2 x 20 mL) and the organic layer was then dried ($MgSO_4$), filtered and concentrated *in vacuo*. The resulting solid was recrystallised from petrol:ethyl acetate to afford phenyl 2,3,4,6-acteyl-1-thio-β-D-glucopyranoside **5.20** (4.80 g, 82%) as a colourless crystalline solid; m.p. 114-116 °C (ethyl acetate) [lit. 113-114 °C (ethyl acetate)][36]; $[\alpha]_D^{20}$ -12 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ -16.5 (*c*, 1.0 in $CHCl_3$)][34]; δ_H (400 MHz, $CDCl_3$) 1.99 (3H, s, C(O)C$\underline{H}_3$), 2.01 (3H, s, C(O)C$\underline{H}_3$), 2.08 (3H, s, C(O)C$\underline{H}_3$), 2.08 (3H, s, C(O)C$\underline{H}_3$), 3.72 (1H, m, H-5), 4.13 - 4.25 (2H, m, H-6 & H-6'), 4.71 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 4.97 (1H, t, *J* 9.8 Hz, H-2), 5.04 (1H, t, *J* 9.8 Hz, H-4), 5.22 (1H, t, *J* 9.0 Hz, H-3), 7.25 - 7.34 (3H, m, Ar-H x 3), 7.46 - 7.52 (2H, m, Ar-H x 2)[36]; δ_C (100.5 MHz, $CDCl_3$) 20.5 (q, C(O)$\underline{C}H_3$), 20.6 (q, C(O)$\underline{C}H_3$), 20.7 (q, C(O)$\underline{C}H_3$), 20.7 (q, C(O)$\underline{C}H_3$), 62.1 (t, C-6), 68.2 (d, C-4), 69.9 (d, C-2), 73.9 (C-3), 75.8 (C-5), 85.7 (C-1), 128.4, 128.9, 131.6, 133.1 (Ar-C x 4), 169.2 (q, C(O)CH_3), 169.4 (q, C(O)CH_3), 170.1 (q, C(O)CH_3), 170.5 (q, C(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{20}H_{24}O_9SNa^+$: 463.1033. Found: 463.1039 (MNa^+).

_p_-Methoxyphenyl 2,3,4,6-acetyl-1-thio-β-D-glucopyranoside 5.21[37]



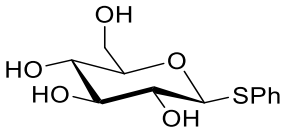
1,2,3,4,6-Penta-*O*-acetyl-β-D-glucopyranose **5.18** (5.86 g, 15 mmol) was dissolved in CH_2Cl_2 (35 mL) and the resulting solution was cooled to 0 °C. *p*-Methoxythiophenol (3.38 mL, 27.5 mmol) was added to the reaction, followed by dropwise addition of $BF_3.OEt_2$ (3.39 mL, 27.5 mmol) with constant stirring. The reaction was warmed to room temperature, and after 2 h t.l.c. (petrol:Et_2O 1:1) indicated complete consumption of starting material (R_f 0.5) and formation of a single product (R_f 0.6). The reaction mixture was diluted with CH_2Cl_2 (40 mL), washed with $NaHCO_3$ (sat. aqueous soln., 3 x 40 mL), followed by brine (2 x 20 mL) and the organic layer was then dried ($MgSO_4$), filtered and concentrated *in vacuo*. The resulting solid was pre-absorbed onto silica and purified by flash column chromatography (petrol:ethyl acetate 2:1) to afford p-methoxyphenyl 2,3,4,6-acteyl-1-thio-β-D-glucopyranoside **5.21** (6.83 g, 97%) as a colourless solid; m.p. 95-96 °C (ethyl acetate) [lit. 92-93 °C (ethyl acetate)][28]; $[\alpha]_D^{20}$ -22 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ -24.3 (*c*, 2.0 in $CHCl_3$)][28]; δ_H (400 MHz, $CDCl_3$)[28] 1.97 (3H, s, C(O)C$\underline{H}_3$), 2.00 (3H, s, C(O)C$\underline{H}_3$), 2.07 (3H, s, C(O)C$\underline{H}_3$), 2.09 (3H, s, C(O)C$\underline{H}_3$), 3.67 (1H, dt, $J_{4,5}$ 10.1 Hz, $J_{5,6}$ 3.6 Hz, $J_{5,6'}$ 3.6 Hz, H-5), 3.81 (3H, s, OC$\underline{H}_3$), 4.16 - 4.20 (2H, m, H-6 & H-6'), 4.55 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 4.88 (1H, t, *J* 9.6 Hz, H-2), 4.98 (1H, t, *J* 9.6 Hz, H-4), 5.19 (1H, t, *J* 9.4 Hz, H-3), 6.84 (2H, d, *J* 8.6 Hz, Ar-H x 2), 7.43 (2H, d, *J* 8.6 Hz, Ar-H x 2); δ_C (100.5 MHz, $CDCl_3$)[28] 20.5 (q, C(O)$\underline{C}H_3$), 20.6 (q, C(O)$\underline{C}H_3$), 20.7 (q, C(O)$\underline{C}H_3$), 20.8 (q, C(O)$\underline{C}H_3$), 55.3 (q, O$\underline{C}H_3$), 62.1 (t, C-6), 68.2 (d, C-4), 69.9 (d, C-2), 74.0 (d, C-3), 75.7 (d, C-5), 85.6 (d, C-1), 114.4 (d, Ar-C x 2), 120.8 (s, Ar-C), 136.5 (D, Ar-C x 2), 160.4 (s, Ar-C), 169.2 (s, $\underline{C}(O)CH_3$), 169.4 (s, $\underline{C}(O)CH_3$), 170.2 (s, $\underline{C}(O)CH_3$), 170.5 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{21}H_{26}O_{10}SNa^+$: 493.1139. Found: 493.1142 (MNa^+).

Ethyl 1-thio-β-D-glucopyranoside 5.22[38]



NaOMe (0.32 g, 6 mmol) was added to a stirred solution of ethyl 2,3,4,6-acteyl-1-thio-β-D-glucopyranoside **5.19** (3.53 g, 9 mmol) in MeOH (45 mL). After 1 h t.l.c. (ethyl acetate:MeOH 9:1) indicated complete consumption of starting material (R_f 0.9) and formation of a single product (R_f 0.3). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until the pH was neutral. The mixture was then filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (ethyl acetate:MeOH 9:1) to afford ethyl 1-thio-β-D-glucopyranoside **5.22** (1.73 g, 86%) as a colourless oil; $[\alpha]_D^{20}$ -54 (*c*, 1.0 in water) [lit. $[\alpha]_D^{20}$ -55.1 (*c*, 1.0 in water)][39]; δ_H (400 MHz, D_2O)[38] 1.17 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.56 - 2.74 (2H, m, $SC\underline{H}_2CH_3$), 3.21 (1H, t, *J* 9.8 Hz, H-2), 3.26 - 3.42 (3H, m, H-3, H-4 & H-5), 3.59 (1H, dd, $J_{5,6}$ 5.5 Hz, $J_{6,6'}$ 12.5 Hz, H-6), 3.79 (1H, d, $J_{6,6'}$ 12.1 Hz, H-6), 4.43 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O)[38] 14.5 (q, $SCH_2\underline{C}H_3$), 24.1 (t, $S\underline{C}H_2CH_3$), 60.9 (t, C-6), 69.5 (d, C-4), 72.2 (d, C-2), 77.2 (d, C-3), 79.8 (d, C-5), 85.0 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_5SNa^+$: 247.0611. Found: 247.0617 (MNa^+).

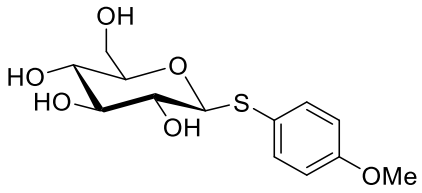
Phenyl-1-thio-β-D-glucopyranoside 5.23[35]



NaOMe (0.36 g, 6.67 mmol) was added to a stirred solution of phenyl 2,3,4,6-acteyl-1-thio-β-D-glucopyranoside **5.20** (4.40 g, 10 mmol) in MeOH (50 mL). After 1 h t.l.c. (ethyl acetate:MeOH 9:1) indicated complete consumption of starting material (R_f 0.9) and formation of a single product (R_f 0.3). Amberlite® IR-120 (H^+) resin was added and the mixture stirred

for 15 min until the pH was neutral. The mixture was then filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (ethyl acetate:MeOH 9:1) to afford phenyl-1-thio-β-D-glucopyranoside **5.23** (2.35 g, 86%) as a white solid; m.p. 128-129 °C (MeOH) [lit. 133 °C (MeOH)][40]; $[\alpha]_D^{20}$ -65.3 (*c*, 1.0 in water) [lit. $[\alpha]_D^{20}$ -70.3 (*c*, 1.0 in water)][40]; δ_H (400 MHz, D_2O)[41] 3.24 (1H, t, *J* 9.0 Hz, H-2), 3.27 - 3.45 (3H, m, H-3, H-4 & H-5), 3.60 (1H, dd, $J_{5,6}$ 5.9 Hz, $J_{6,6'}$ 12.5 Hz, H-6), 3.75 - 3.81 (1H, m, H-6'), 4.68 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 7.24 - 7.34 (3H, m, Ar-H x 3), 7.47 (2H, d, *J* 7.8 Hz, Ar-H x 2); δ_C (100.5 MHz, D_2O)[41] 60.7 (t, C-6), 69.3 (d, C-4), 71.7 (d, C-2), 77.2 (d, C-3), 79.8 (d, C-5), 87.2 (d, C-1), 128.1, 129.3, 131.6, 131.9 (4 x Ar-C); HRMS (ESI-TOF): calcd. for $C_{12}H_{16}O_5SNa^+$: 295.0611. Found: 295.0611 (MNa^+).

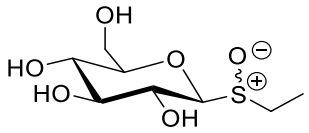
_p_-Methoxyphenyl-1-thio-β-D-glucopyranoside 5.24



NaOMe (0.47 g, 8.67 mmol) was added to a stirred solution of p-methoxyphenyl 2,3,4,6-acteyl-1-thio-β-D-glucopyranoside **5.21** (6.11 g, 13 mmol) in MeOH (65 mL). After 1 h t.l.c. (ethyl acetate:MeOH 9:1) indicated complete consumption of starting material (R_f 0.9) and formation of a single product (R_f 0.3). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until the pH was neutral. The mixture was then filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (ethyl acetate:MeOH 9:1) to afford *p*-methoxyphenyl-1-thio-β-D-glucopyranoside **5.24** (3.22 g, 82%) as a colourless foam; $[\alpha]_D^{20}$ -33 (*c*, 1.0 in MeOH) [lit. $[\alpha]_D^{20}$ -41.0 (*c*, 1.0 in MeOH)][42]; δ_H (400 MHz, D_2O)[43] 3.14 (1H, t, *J* 9.4 Hz, H-2), 3.21 - 3.33 (2H, m, H-4 & H-5), 3.38 (1H, t, *J* 9.0 Hz, H-3), 3.59 (1H, dd, $J_{6,6'}$ 12.5 Hz, $J_{5,6}$ 5.5 Hz, H-6), 3.71 (3H, s, OC$\underline{H}_3$), 3.73 - 3.79

(1H, m, H-6'), 4.49 (1H, d, $J_{1,2}$ 9.8 Hz, H-1), 6.87 (2H, d, *J* 9.0 Hz, Ar-H x 2), 7.43 (2H, d, *J* 9.0 Hz, Ar-H x 2); δ_C (100.5 MHz, D_2O)[43] 55.4 (q, O<u>C</u>H_3), 60.8 (t, C-6), 69.3 (d, H-4), 71.5 (d, H-2), 77.2 (d, H-3), 79.8 (d, H-5), 87.7 (d, H-1), 114.8 (d, Ar-C x 2), 122.0 (s, Ar-C), 135.1 (d, Ar-C x 2), 159.4 (s, Ar-C); HRMS (ESI-TOF): calcd. for $C_{13}H_{18}O_6SNa^+$: 325.0716. Found: 325.0711 (MNa^+).

Ethyl 1-sulfinyl-β-D-glucopyranoside 5.25



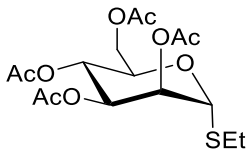
Method 1

Ethyl 1-thio-β-D-glucopyranoside **5.22** (0.672 g, 3 mmol) was dissolved in MeOH (8.1 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0417 g, 0.15 mmol) was added followed by H_2O_2 (0.21 mL, 50% w/w in water, 3.6 mmol), and the reaction was then stirred for 6 h. The reaction mixture was concentrated *in vacuo* and the residue produced was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-sulfinyl-β-D-glucopyranoside **5.25** (0.551 g, 76%) as an orange oil (1:1 mixture of diastereoisomers); υ_{max} (neat) 1042 cm^{-1} (S=O); δ_H (400 MHz, D_2O) 1.17 - 1.28 (6H, m, $SCH_2C\underline{H}_{3a}$ & $SCH_2C\underline{H}_{3b}$), 2.81 - 2.94 (2H, m, $SC\underline{H}H'CH_{3a}$ & $SC\underline{H}H'CH_{3b}$), 3.04 - 3.19 (2H, m, $SCH\underline{H}'CH_{3a}$ & $SCH\underline{H}'CH_{3b}$), 3.29 - 3.73 (10H, m, $H\text{-}2_a$, $H\text{-}2_b$, $H\text{-}3_a$, $H\text{-}3_b$, $H\text{-}4_a$, $H\text{-}4_b$, $H\text{-}5_a$, $H\text{-}5_b$, $H\text{-}6_a$ & $H\text{-}6_b$), 3.83 (2H, dd, $J_{5,6'}$ 6.5 Hz, $J_{6,6'}$ 11.9 Hz, $H\text{-}6'_a$ & $H\text{-}6'_b$), 4.17 (1H, d, $J_{1,2}$ 9.8 Hz, $H\text{-}1_b$), 4.47 (1H, d, $J_{1,2}$ 9.8 Hz, $H\text{-}1_a$); δ_C (100.5 MHz, D_2O) 6.5 (q, $SCH_2\underline{C}H_{3a}$ & $SCH_2\underline{C}H_{3b}$), 39.4 (t, $S\underline{C}H_2CH_{3a}$), 40.1 (t, $S\underline{C}H_2CH_{3b}$), 60.4 (t, $C\text{-}6_b$), 60.7 (t, $C\text{-}6_a$), 67.7 (d, $C\text{-}2_b$), 68.6 (d, $C\text{-}4_b$), 68.8 (d, $C\text{-}4_a$), 69.1 (d, $C\text{-}2_a$), 76.9 (d, $C\text{-}3_b$), 77.1 (d, $C\text{-}3_a$), 80.2 (d, $C\text{-}5_b$), 80.7 (d, $C\text{-}5_a$), 87.7 (d, $C\text{-}1_b$), 90.4 (d, $C\text{-}1_a$); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0563 (MNa^+).

Method 2

Ethyl 1-thio-β-D-glucopyranoside **5.22** (0.672 g, 3 mmol) was dissolved in water (8.1 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0417 g, 0.15 mmol) was added followed by H_2O_2 (0.21 mL, 50% w/w in water, 3.6 mmol), and the reaction was then stirred for 6 h. The reaction mixture was freeze-dried and the residue produced was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-sulfinyl-β-D-glucopyranoside **5.25** (0.512 g, 71%) as an orange oil (1:1 mixture of diastereoisomers), identical in all respects to that produced by *Method 1*.

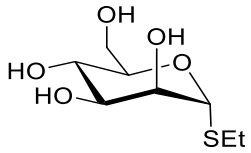
Ethyl 2,3,4,6-acetyl-1-thio-α-D-mannopyranoside 5.29[44]



1,2,3,4,6-Penta-*O*-acetyl-α-D-mannopyranose **5.28** (11.71 g, 30 mmol) was dissolved in CH_2Cl_2 (70 mL). Ethanethiol (3.96 mL, 55 mmol) was added, followed by the dropwise addition of $BF_3.OEt_2$ (6.77 mL, 55 mmol) with constant stirring. The reaction was stirred at room temperature, and after 24 h t.l.c. (petrol:ethyl acetate 2:1) indicated complete consumption of starting material (R_f 0.2) and formation of a single product (R_f 0.3). The reaction mixture was quenched with $NaHCO_3$ (sat. aqueous soln., 2 x 100 mL), and then washed with brine (80 mL) and the organic layer was then dried ($MgSO_4$), filtered and concentrated *in vacuo*. The solid was crystallised from CH_2Cl_2:petrol, and the mother liquors concentrated *in vacuo* to afford ethyl 2,3,4,6-acetyl-1-thio-α-D-mannopyranoside **5.29** (3.65 g, 31%) as a colourless crystalline solid; m.p. 109-110 °C (CH_2Cl_2) [lit. 106-108 °C (EtOH)][34]; $[\alpha]_D^{20}$ +94 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ +118.7 (*c*, 1.0 in $CHCl_3$)][34]; δ_H (400 MHz, $CDCl_3$)[34] 1.30 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 1.99 (3H, s, $C(O)C\underline{H}_3$), 2.05 (3H, s, $C(O)C\underline{H}_3$), 2.09 (3H, s, $C(O)C\underline{H}_3$), 2.16 (3H, s, $C(O)C\underline{H}_3$), 2.64 (2H, m, $SC\underline{H}_2CH_3$), 4.10 (1H, dd, $J_{6,6'}$ 12.1 Hz,

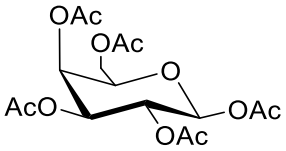
$J_{5,6}$ 1.6 Hz, H-6), 4.31 (1H, dd, $J_{6,6'}$ 12.1 Hz, $J_{5,6'}$ 5.1 Hz, H-6), 4.36 - 4.43 (1H, m, H-5), 5.23 - 5.36 (4H, m, H-1, H-2, H-3 & H-4); δ_C (100.5 MHz, $CDCl_3$)[34] 14.7 (q, $SCH_2\underline{C}H_3$), 20.6 (q, $C(O)\underline{C}H_3$), 20.7 (q, $C(O)\underline{C}H_3$), 20.7 (q, $C(O)\underline{C}H_3$), 20.9 (q, $C(O)\underline{C}H_3$), 25.5 (t, $S\underline{C}H_2CH_3$), 62.4 (t, C-6), 66.4 (d, C-4), 68.9 (d, C-5), 69.5 (d, C-3), 71.2 (d, C-2), 82.3 (d, C-1), 169.7 (s, $\underline{C}(O)CH_3$), 169.7 (s, $\underline{C}(O)CH_3$), 170.0 (s, $\underline{C}(O)CH_3$), 170.6 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{16}H_{24}O_9SNa^+$: 415.1033. Found: 415.1051 (MNa^+).

Ethyl 1-thio-α-D-mannopyranoside 5.30[45]



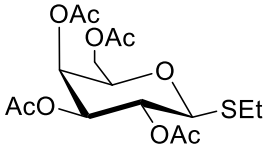
NaOMe (0.32 g, 6 mmol) was added to a stirred solution of ethyl 2,3,4,6-acetyl-1-thio-α-D-mannopyranoside **5.29** (3.53 g, 9 mmol) in MeOH (45 mL). After 1 h t.l.c. (CH_2Cl_2:MeOH 7:1) indicated complete consumption of starting material (R_f 0.8) and formation of a single product (R_f 0.1). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until neutral pH. The mixture was then filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-thio-α-D-mannopyranoside **5.30** (1.83 g, 91%) as a white solid; m.p. 105-106 °C (MeOH) [lit. 126-127 °C (EtOH)][46]; $[\alpha]_D^{20}$ +177.9 (*c*, 1.0 in water) [lit. $[\alpha]_D^{20}$ +20.2 (*c*, 0.85 in water)][46]; δ_H (400 MHz, D_2O)[45] 1.16 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.47 - 2.67 (2H, m, $SC\underline{H}_2CH_3$), 3.55 (1H, t, *J* 9.8 Hz, H-4), 3.62 - 3.69 (2H, m, H-3 & H-6), 3.73 - 3.79 (1H, m, H-6'), 3.84 - 3.90 (1H, m, H-5), 3.91 - 3.93 (1H, m, H-2), 5.20 (1H, s, H-1); δ_C (100.5 MHz, D_2O) 14.1 (q, $SCH_2\underline{C}H_3$), 24.8 (t, $S\underline{C}H_2CH_3$), 60.8 (t, C-6), 67.1 (d, C-4), 71.1 (d, C-3), 71.8 (d, C-2), 73.0 (d, C-5), 84.2 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_5SNa^+$: 247.0611. Found: 247.0606 (MNa^+).

1,2,3,4,6-Penta-*O*-acetyl-β-D-galactopyranose 5.32[47]



Sodium acetate (45.9 g, 560 mmol) was suspended in acetic anhydride (320 mL). The resulting solution was stirred and heated to 120 °C. D-Galactopyranose **5.31** (50.44 g, 280 mmol) was added over 100 min. After a further 20 min the reaction mixture was cooled to room temperature, and then poured onto ice (400 mL). The resulting white precipitate was filtered off and recrystallised from EtOH to afford 1,2,3,4,6-penta-*O*-acetyl-β-D-galactopyranose **5.32** (52.15 g, 48%); m.p. 139-140 °C (EtOH) [lit. 143-144 °C (EtOH)][48]; $[\alpha]_D^{20}$ +23 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ +27.4 (*c*, 1.0 in $CHCl_3$)][48]; δ_H (400 MHz, $CDCl_3$)[47] 1.98 (3H, s, C(O)C$\underline{H}_3$), 2.03 (6H, s, 2 x C(O)C$\underline{H}_3$), 2.11 (3H, s, C(O)C$\underline{H}_3$), 2.15 (3H, s, C(O)C$\underline{H}_3$), 4.04 (1H, t, *J* 6.7 Hz, H-5), 4.08 - 4.18 (2H, m, H-6 & H-6'), 5.07 (1H, dd, $J_{2,3}$ 10.4 Hz, $J_{3,4}$ 3.3 Hz, H-3), 5.32 (1H, dd, $J_{2,3}$ 10.2 Hz, $J_{1,2}$ 8.6 Hz, H-2), 5.41 (1H, d, $J_{3,4}$ 3.5 Hz, H-4), 5.69 (1H, d, $J_{1,2}$ 8.2 Hz, H-1); δ_C (100.5 MHz, $CDCl_3$)[47] 20.5 (q, C(O)$\underline{C}H_3$), 20.6 (q, C(O)$\underline{C}H_3$), 20.6 (q, 2 x C(O)$\underline{C}H_3$), 20.8 (q, C(O)$\underline{C}H_3$), 61.0 (t, C-6), 66.8 (d, C-4), 67.8 (d, C-2), 70.8 (d, C-3), 71.7 (d, C-5), 92.1 (d, C-1), 168.9 (s, $\underline{C}$(O)CH_3), 169.3 (s, $\underline{C}$(O)CH_3), 169.9 (s, $\underline{C}$(O)CH_3), 170.1 (s, $\underline{C}$(O)CH_3), 170.3 (s, $\underline{C}$(O)CH_3); HRMS (ESI-TOF): calcd. for $C_{16}H_{22}O_{11}Na^+$: 413.1054. Found: 413.1071 (MNa^+).

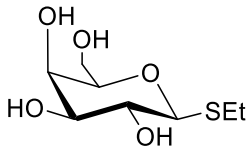
Ethyl 2,3,4,6-acetyl-1-thio-β-D-galactopyranoside 5.33[49]



1,2,3,4,6-Penta-*O*-acetyl-β-D-galactopyranose **5.32** (5.86 g, 15 mmol) was dissolved in CH_2Cl_2 (35 mL) and the resulting solution cooled to 0 °C. Ethanethiol (1.98 mL, 27.5 mmol) was

added, followed by the dropwise addition of $BF_3.OEt_2$ (3.39 mL, 27.5 mmol) with constant stirring. The reaction was warmed to room temperature, and after 2 h t.l.c. (petrol:Et_2O 2:1) indicated complete consumption of starting material (R_f 0.4) and the formation of a single product (R_f 0.5). The reaction mixture was diluted with CH_2Cl_2 (40 mL), washed with aqueous $NaHCO_3$ (sat. aqueous soln., 3 x 40 mL), followed by brine (2 x 20 mL) and the organic layer was then dried ($MgSO_4$), filtered and concentrated *in vacuo*. The resulting solid was recrystallised from the minimum amount of hot EtOH to afford ethyl 2,3,4,6-acetyl-1-thio-β-D-galactopyranoside **5.33** (4.6 g, 78%) as a white solid; m.p. 74-76 °C (EtOH) [lit. 73-75 °C (Et_2O)][34]; $[\alpha]_D^{20}$ -4 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ -9.2 (*c*, 1.0 in $CHCl_3$)][34]; δ_H (400 MHz, $CDCl_3$)[34] 1.28 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 1.98 (3H, s, $C(O)C\underline{H}_3$), 2.04 (3H, s, $C(O)C\underline{H}_3$), 2.06 (3H, s, $C(O)C\underline{H}_3$), 2.15 (3H, s, $C(O)C\underline{H}_3$), 2.64 - 2.81 (2H, m, $SC\underline{H}_2CH_3$), 3.93 (1H, dd, $J_{5,6}$ 7.4 Hz, $J_{5,6'}$ 5.9 Hz, H-5), 4.07 - 4.20 (2H, m, H-6 & H-6'), 4.49 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 5.04 (1H, dd, $J_{2,3}$ 10.0 Hz, $J_{3,4}$ 3.3 Hz, H-3), 5.23 (1H, t, *J* 9.8 Hz, H-2), 5.42 (1H, d, $J_{3,4}$ 3.1 Hz, H-4); δ_C (100.5 MHz, $CDCl_3$)[34] 14.8 (q, $SCH_2\underline{C}H_3$), 20.6 (q, $C(O)\underline{C}H_3$), 20.6 (q, 2 x $C(O)\underline{C}H_3$), 20.8 (q, $C(O)\underline{C}H_3$), 24.3 ($S\underline{C}H_2CH_3$), 61.5 (t, C-6), 67.2 (d, C-2), 67.3 (d, C-4), 71.9 (d, C-3), 74.4 (d, C-5), 84.1 (d, C-1), 169.5 (s, $\underline{C}(O)CH_3$), 170.0 (s, $\underline{C}(O)CH_3$), 170.2 (s, $\underline{C}(O)CH_3$), 170.4 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{16}H_{24}O_9SNa^+$: 415.1033. Found: 415.1048 (MNa^+).

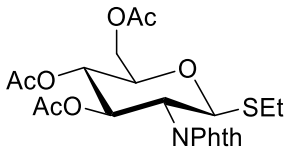
Ethyl 1-thio-β-D-galactopyranoside 5.34[50]



NaOMe (0.27 g, 5 mmol) was added to a stirred solution of ethyl 2,3,4,6-acetyl-1-thio-α-D-galactopyranoside **5.33** (2.94 g, 7.5 mmol) in MeOH (40 mL). After 1 h t.l.c. (CH_2Cl_2:MeOH 7:1) indicated complete consumption of starting material (R_f 0.7) and the formation of a single

product (R_f 0.2). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until the pH was neutral. The mixture was then filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-thio-α-D-galactopyranoside **5.34** (1.46 g, 87%) as a white solid; m.p. 124-125 °C (MeOH) [lit. 122-123 °C (MeOH)][50]; $[\alpha]_D^{20}$ -15 (*c*, 1.0 in water) [lit. $[\alpha]_D^{20}$ -23.5 (*c*, 1.1 in water)][50]; δ_H (400 MHz, D_2O) 1.16 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.56 - 2.73 (2H, m, $SC\underline{H}_2CH_3$), 3.44 (1H, t, *J* 9.8 Hz, H-2), 3.50 - 3.56 (1H, m, H-3), 3.56 - 3.68 (3H, m, H-5, H-6 & H-6'), 3.86 (1H, s, H-4), 4.36 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O) 14.4 (q, $SCH_2\underline{C}H_3$), 24.2 (t, $S\underline{C}H_2CH_3$), 61.0 (t, C-6), 68.8 (d, C-2), 69.5 (d, C-3), 73.9 (d, C-5), 78.8 (d, C-4), 85.5 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_5SNa^+$: 247.0611. Found: 247.0602 (MNa^+).

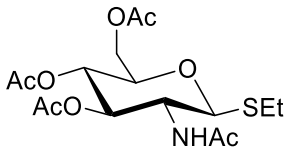
Ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside 5.36[44]



1,3,4,6-Tetra-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside **5.35** (12.0 g, 25 mmol) was dissolved in CH_2Cl_2 (100 mL). Ethanethiol (2.16 mL, 37.5 mmol) was added, followed by dropwise addition of $BF_3.OEt_2$ (4.62 mL, 37.5 mmol). The reaction was heated to reflux under an atmosphere of nitrogen for 3 h, at which point t.l.c. (petrol:ethyl acetate 1:1) indicated complete consumption of starting material (R_f 0.5) and formation of a single product (R_f 0.6). The reaction mixture was cooled to room temperature, diluted with CH_2Cl_2 (80 mL), quenched with aqueous $NaHCO_3$ (sat. aqueous soln., 2 x 80 mL), washed with brine (40 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was pre-absorbed onto silica and purified by flash column chromatography (petrol:ethyl acetate 3:2) to afford ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside **5.36** (10.24 g, 85%) as a colourless

solid; m.p. 116-118 °C (ethyl acetate) [lit. 118-119 °C (ethyl acetate)][34]; $[\alpha]_D^{20}$ +42 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ +46.5 (*c*, 1.0 in $CHCl_3$)][34]; δ_H (400 MHz, $CDCl_3$)[34] 1.21 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 1.85 (3H, s, $C(O)C\underline{H}_3$), 2.03 (3H, s, $C(O)C\underline{H}_3$), 2.10 (3H, s, $C(O)C\underline{H}_3$), 2.58 - 2.76 (2H, m, $SC\underline{H}_2CH_3$), 3.86 - 3.92 (1H, m, H-5), 4.17 (1H, dd, $J_{6,6'}$ 11.7 Hz, $J_{5,6}$ 1.6 Hz, H-6), 4.30 (1H, dd, $J_{6,6'}$ 12.5 Hz, $J_{5,6'}$ 4.7 Hz, H-6'), 4.39 (1H, t, *J* 10.4 Hz, H-2), 5.17 (1H, t, *J* 9.6 Hz, H-4), 5.48 (1H, d, $J_{1,2}$ 10.6 Hz, H-1), 5.82 (1H, t, *J* 9.6 Hz, H-3), 7.71 - 7.77 (2H, m, 2 x Ar-H), 7.83 - 7.89 (2H, m, 2 x Ar-H); δ_C (100.5 MHz, $CDCl_3$)[34] 14.9 (q, $SCH_2\underline{C}H_3$), 20.4 (q, $C(O)\underline{C}H_3$), 20.6 (q, $C(O)\underline{C}H_3$), 20.8 (q, $C(O)\underline{C}H_3$), 24.3 (t, $S\underline{C}H_2CH_3$), 53.7 (d, H-2), 62.3 (t, C-6), 68.9 (d, C-4), 71.5 (d, C-3), 75.9 (d, C-5), 81.2 (d, H-1), 123.7 (d, 2 x Ar-C), 134.3 (d, 2 x Ar-C), 134.4 (s, 2 x Ar-C), 169.4 (s, $\underline{C}(O)CH_3$), 170.1 (s, $\underline{C}(O)CH_3$), 170.7 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{22}H_{25}NO_9SNa^+$: 502.1142. Found: 502.1164 (MNa^+).

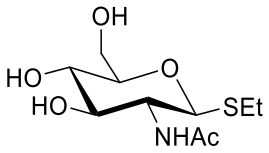
Ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-1-thio-β-D-glucopyranoside 5.37[51]



Ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside **5.36** (9.59 g, 20 mmol) was dissolved in EtOH (100 mL) and $H_2NNH_2.H_2O$ (30.0 mL, 800 mmol) was added. The reaction was then refluxed for 16 h, at which point t.l.c. (petrol:ethyl acetate 1:1) indicated complete consumption of starting material (R_f 0.6) and the formation of a single product (R_f 0). The reaction mixture was then cooled, concentrated *in vacuo* and then co-evaporated with toluene (3 x 100 mL). The residue was then dissolved in pyridine (80 mL) and acetic anhydride (40 mL) was added. After 16 h the reaction mixture was concentrated *in vacuo*, and the residue was dissolved in ethyl acetate. The solution was then washed with aqueous HCl (1M, 100 mL), filtered, washed with water (100 mL) and aqueous $NaHCO_3$ (sat. aqueous soln., 100 mL). The organic phase was then dried ($MgSO_4$), filtered and concentrated

in vacuo. The residue was then recrystallised (EtOH:petrol) to afford ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-1-thio-β-D-glucopyranoside **5.37** (3.25 g, 42%) as a white crystalline solid; m.p. 175-176 °C (EtOH) [lit. 194-195 °C (EtOH)][38]; $[\alpha]_D^{20}$ -48 (*c*, 1.0 in $CHCl_3$) [lit. $[\alpha]_D^{20}$ -52 (*c*, 1.0 in $CHCl_3$)][52]; δ_H (400 MHz, $CDCl_3$)[38] 1.27 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 1.96 (3H, s, $C(O)C\underline{H}_3$), 2.02 (3H, s, $C(O)C\underline{H}_3$), 2.03 (3H, s, $C(O)C\underline{H}_3$), 2.08 (3H, s, $C(O)C\underline{H}_3$), 2.65 - 2.79 (2H, m, $SC\underline{H}_2CH_3$), 3.66 - 3.72 (1H, m, H-5), 4.05 - 4.16 (2H, m, H-2 & H-6), 4.23 (1H, dd, $J_{6,6'}$ 12.1 Hz, $J_{5,6'}$ 4.7 Hz, H-6'), 4.59 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 5.09 (1H, t, *J* 9.8 Hz, H-4), 5.17 (1H, t, *J* 9.8 Hz, H-3), 5.51 (1H, d, $J_{N\underline{H},2}$ 9.4 Hz, N<u>H</u>); δ_C (100.5 MHz, $CDCl_3$)[48] 14.7 (q, $SCH_2\underline{C}H_3$), 20.6 (q, $C(O)\underline{C}H_3$), 20.7 (q, $C(O)\underline{C}H_3$), 20.7 (q, $C(O)\underline{C}H_3$), 23.3 (q, $C(O)\underline{C}H_3$), 24.2 (t, $S\underline{C}H_2CH_3$), 53.3 (d, C-2), 62.3 (t, C-6), 68.4 (d, C-4), 73.9 (d, C-3), 75.9 (d, C-5), 84.4 (d, C-1), 169.3 (s, $\underline{C}(O)CH_3$), 170.0 (s, $\underline{C}(O)CH_3$), 170.7 (s, $\underline{C}(O)CH_3$), 171.1 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{16}H_{25}NO_8SNa^+$: 414.1193. Found: 414.1199 (MNa^+).

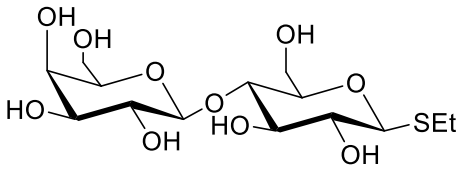
Ethyl 2-deoxy-2-acetamido-1-thio-β-D-glucopyranoside 5.38



NaOMe (0.22 g, 4 mmol) was added to a stirred solution of ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-acetamido-1-thio-β-D-glucopyranoside **5.37** (3.13 g, 8 mmol) in MeOH (40 mL). After 1 h t.l.c. (CH_2Cl_2:MeOH 7:1) indicated complete consumption of starting material (R_f 0.7) and the formation of a single product (R_f 0.2). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until the pH was neutral. The mixture was then filtered and concentrated *in vacuo*. The resulting residue was pre-absorbed onto silica and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 2-deoxy-2-acetamido-1-thio-β-D-glucopyranoside **5.38** (1.26 g, 59%) as a white solid; m.p. 118-120 °C (MeOH) [lit. 215 °C (MeOH)][53]; $[\alpha]_D^{20}$ -43 (*c*, 1.0 in water) [lit. $[\alpha]_D^{20}$ -55 (*c*, 1.0 in water)][53]; δ_H (400 MHz, D_2O)[38]

1.12 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 1.92 (3H, s, $C(O)C\underline{H}_3$), 2.51 - 2.71 (2H, m, $SC\underline{H}_2CH_3$), 3.31 - 3.38 (2H, m, H-4 & H-5), 3.39 - 3.47 (1H, m, H-3), 3.57 - 3.67 (2H, m, H-2 & H-6), 3.79 (1H, d, $J_{6,6'}$ 12.5 Hz, H-6'), 4.51 (1H, d, $J_{1,2}$ 10.6 Hz, H-1); δ_C (100.5 MHz, D_2O)[38] 14.2 (q, $SCH_2\underline{C}H_3$), 22.1 (q, $C(O)\underline{C}H_3$), 24.4 (t, $S\underline{C}H_2CH_3$), 54.7 (d, C-2), 60.8 (t, C-6), 69.8 (d, C-4), 75.1 (d, C-3), 79.8 (d, C-5), 83.9 (d, C-1), 174.4 (s, $\underline{C}(O)CH_3$); HRMS (ESI-TOF): calcd. for $C_{10}H_{19}NO_5SNa^+$: 288.0876. Found: 288.0878 (MNa^+).

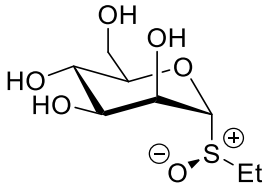
Ethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside 5.40[54]



2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→4)-1,2,3,6-tetra-*O*-acetyl-β-D-glucopyranoside **5.39** (16.96 g, 25 mmol) was dissolved in CH_2Cl_2 (100 mL) and the resulting solution cooled to 0 °C. Ethanethiol (2.79 mL, 37.5 mmol) was added, followed by the dropwise addition of $BF_3.OEt_2$ (4.62 mL, 37.5 mmol) with constant stirring. The reaction was warmed to room temperature, and after 2 h t.l.c. (petrol:Et_2O 2:1) indicated complete consumption of starting material (R_f 0.3) and formation of a single product (R_f 0.4). The reaction mixture was diluted with CH_2Cl_2 (100 mL), washed with $NaHCO_3$ (sat. aqueous soln., 3 x 100 mL), followed by brine (2 x 50 mL) and the organic layer was then dried ($MgSO_4$), filtered and concentrated *in vacuo*. The resulting solid was then dissolved in MeOH (70 mL) and NaOMe (0.64 g, 11.9 mmol) was added. After 1 h, t.l.c (CH_2Cl_2:MeOH 7:1) indicated complete consumption of starting material (R_f 0.6) and the formation of a single product (R_f 0.2). Amberlite® IR-120 (H^+) resin was added and the mixture stirred for 15 min until neutral pH. The mixture was then filtered and concentrated *in vacuo*. The resulting solid was recrystallised from the minimum amount of hot EtOH to afford ethyl β-D-galactopyranosyl-

(1→4)-1-thio-β-D-glucopyranoside **5.40** (4.06 g, 42%) as a white crystalline solid; m.p. 185-187 °C (EtOH) [lit. 189-192 °C (EtOH)][54]; $[\alpha]_D^{20}$ -8 (*c*, 1.0 in water) [lit. $[\alpha]_D^{20}$ -27.4 (*c*, 0.5 in MeOH)][54]; δ_H (400 MHz, D_2O)[54] 1.17 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.57 - 2.73 (2H, m, $SC\underline{H}_2CH_3$), 3.26 (1H, t, *J* 9.2 Hz, H-2_a), 3.40 - 3.46 (1H, m, H-2_b), 3.49 - 3.71 (8H, m, H-3_a, H-3_b, H-4_a, H-5_a, H-5_b, H-6_a, H-6_b & H-$6'_b$), 3.81 (1H, d, $J_{3,4}$ 3.5 Hz, H-4_b), 3.86 (1H, d, $J_{6,6'}$ 12.5 Hz, H-$6'_a$), 4.34 (1H, d, *J* 7.8 Hz, H-1_b), 4.46 (1H, d, *J* 9.8 Hz, H-1_a); δ_C (100.5 MHz, D_2O) 14.5 (q, $SCH_2\underline{C}H_3$), 24.1 (t, $S\underline{C}H_2CH_3$), 60.2 (t, C-6_a), 61.0 (t, C-6_b), 68.5 (d, C-4_b), 70.9 (d, C-2_b), 71.9 (d, C-2_a), 72.5 (d, C-3_a), 75.3 (d, C-5_a), 75.7 (d, C-3_b), 78.2 (d, C-4_a), 78.6 (d, C-5_b), 84.9 (d, C-1_a), 102.8 (d, C-1_b); HRMS (ESI-TOF): calcd. for $C_{14}H_{26}O_{10}SNa^+$: 409.1139. Found: 409.1135 (MNa^+).

Ethyl 1-(*R*)-sulfinyl-α-D-mannopyranoside 5.41



Method 1

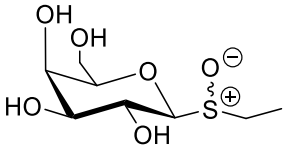
Ethyl 1-thio-α-D-mannopyranoside **5.30** (0.858 g, 3.83 mmol) was dissolved in MeOH (10.3 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0532 g, 0.19 mmol) was added followed by H_2O_2 (0.27 mL, 50% w/w in water, 4.6 mmol), and the reaction was then stirred for 6 h. The reaction mixture was concentrated *in vacuo* and the residue was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-(*R*)-sulfinyl-α-D-mannopyranoside **5.41** (0.645 g, 70%) as a pale orange oil; υ_{max} (neat) 1036 cm^{-1} (S=O); δ_H (400 MHz, D_2O) 1.26 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.75 – 2.86 (1H, m, $SC\underline{H}H'CH_3$), 3.02 - 3.14 (1H, m, $SCH\underline{H}'CH_3$), 3.47 - 3.55 (1H, m, H-5), 3.60 (1H, dd, $J_{6,6'}$ 12.5 Hz, $J_{5,6}$ 5.9 Hz, H-6), 3.67 (1H, t, *J* 9.8 Hz, H-4), 3.73 - 3.79 (1H, dd, $J_{5,6'}$ 2.0 Hz,

$J_{6,6'}$ 12.5 Hz, H-6'), 3.89 (1H, dd, $J_{3,4}$ 9.8 Hz, $J_{2,3}$ 3.5 Hz, H-3_a & H-3_b), 4.29 - 4.33 (1H, m, H-2), 4.71 (1H, s, H-1); δ_C (100.5 MHz, D_2O) 5.2 (q, $SCH_2\underline{C}H_3$), 42.5 (t, $S\underline{C}H_2CH_3$), 60.8 (t, C-6), 65.7 (d, C-4), 66.7 (d, C-2), 70.4 (d, C-3), 78.7 (d, C-5), 92.5 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0557 (MNa^+).

Method 2

Ethyl 1-thio-α-D-mannopyranoside **5.30** (0.574 g, 2.56 mmol) was dissolved in water (6.9 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0356 g, 0.13 mmol) was added followed by H_2O_2 (0.18 mL, 50% w/w in water, 3.07 mmol), and the reaction was then stirred for 6 h. The reaction mixture was freeze-dried and the residue was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-(*R*)-sulfinyl-α-D-mannopyranoside **5.41** (0.303 g, 49%) as a pale orange oil, identical in all respects to the material produced by *Method 1*.

Ethyl 1-sulfinyl-β-D-galactopyranoside 5.42



Method 1

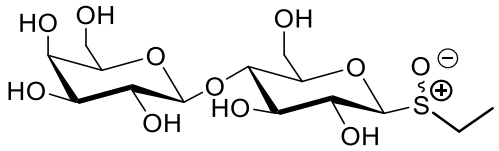
Ethyl 1-thio-β-D-galactopyranoside **5.34** (0.858 g, 3.83 mmol) was dissolved in MeOH (10.3 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0532 g, 0.19 mmol) was added followed by H_2O_2 (0.27 mL, 50% w/w in water, 4.6 mmol), and the reaction was then stirred for 6 h. The reaction mixture was concentrated *in vacuo* and the residue was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-sulfinyl-β-D-galactopyranoside **5.42** (0.611 g, 66%) as a pale orange oil (1:0.7 mixture of diastereoisomers); υ_{max} (neat) 1050 cm^{-1} (S=O); δ_H (400 MHz, D_2O) 1.19 (3H, t, *J* 7.6 Hz,

$SCH_2C\underline{H}_{3b}$), 1.24 (3H, t, *J* 8.2 Hz, $SCH_2C\underline{H}_{3a}$), 2.82 - 2.95 (2H, m, $SC\underline{H}H'CH_{3a}$ & $SC\underline{H}H'CH_{3b}$), 3.03 - 3.17 (2H, m, $SCH\underline{H}'CH_{3a}$ & $SCH\underline{H}'CH_{3b}$), 3.58 - 3.92 (12H, m, $H\text{-}2_a$, $H\text{-}2_b$, $H\text{-}3_a$, $H\text{-}3_b$, $H\text{-}4_a$, $H\text{-}4_b$, $H\text{-}5_a$, $H\text{-}5_b$, $H\text{-}6_a$, $H\text{-}6_b$, $H\text{-}6'_a$ & $H\text{-}6'_b$), 4.08 (1H, d, $J_{1,2}$ 9.8 Hz, $H\text{-}1_b$), 4.40 (1H, d, $J_{1,2}$ 9.8 Hz, $H\text{-}1_a$); δ_C (100.5 MHz, D_2O) 6.4 (q, $SCH_2\underline{C}H_{3a}$), 6.5 (q, $SCH_2\underline{C}H_{3b}$), 39.4 (t, $S\underline{C}H_2CH_{3a}$), 40.2 (t, $S\underline{C}H_2CH_{3b}$), 61.0 (t, $C\text{-}6_a$), 61.2 (t, $C\text{-}6_b$), 65.0 (d, $C\text{-}2_b$), 66.5 (d, $C\text{-}2_a$), 68.6 (d, $C\text{-}3_b$), 68.8 (d, $C\text{-}3_a$), 73.9 (d, $C\text{-}4_b$), 73.9 (d, $C\text{-}4_a$), 79.9 (d, $C\text{-}5_b$), 80.2 (d, $C\text{-}5_a$), 88.4 (d, $C\text{-}1_b$), 91.1 (d, $C\text{-}1_a$); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0554 (MNa^+).

Method 2

Ethyl 1-thio-β-D-galactopyranoside **5.34** (0.672 g, 3 mmol) was dissolved in water (8.1 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0417 g, 0.15 mmol) was added followed by H_2O_2 (0.21 mL, 50% w/w in water, 3.6 mmol), and the reaction was then stirred for 6 h. The reaction mixture was freeze-dried and the residue was pre-absorbed onto Florisil® and purified by flash column chromatography (CH_2Cl_2:MeOH 7:1) to afford ethyl 1-sulfinyl-β-D-galactopyranoside **5.42** (0.456 g, 63%) as a pale orange oil (1:0.9 mixture of diastereoisomers), identical in all respects to the material produced by *method 1*.

Ethyl β-D-galactopyranosyl-(1→4)-1-sulfinyl-β-D-glucopyranoside 5.44



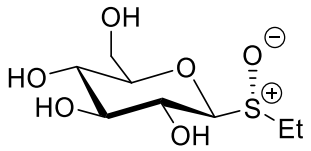
Ethyl β-D-galactopyranosyl-(1→4)-1-thio-β-D-glucopyranoside **5.40** (1.16 g, 3 mmol) was dissolved in MeOH (8.1 mL). 1,10-Ethyleneisoalloxazinium chloride **5.7** (0.0417 g, 0.015 mmol) was added followed by H_2O_2 (0.21 mL, 50% w/w in water, 3.6 mmol), and the reaction was then stirred for 6 h. The reaction mixture was freeze-dried and the residue was

pre-absorbed onto Florisil® and purified by flash column chromatography (MeCN:MeOH 5:1) to afford ethyl β-D-galactopyranosyl-(1→4)-1-sulfinyl-β-D-glucopyranoside **5.44** (0.51 g, 42%) as a colourless solid (1:1 mixture of diastereoisomers); m.p. 175-177 °C (MeOH); υ_{max} (neat) 1056 cm^{-1} (S=O); δ_H (400 MHz, D_2O) 1.17 - 1.29 (6H, m, $SCH_2C\underline{H}_{3a}$ & $SCH_2C\underline{H}_{3b}$), 2.83 - 2.94 (2H, m, $SC\underline{H}H'CH_{3a}$ & $SC\underline{H}H'CH_{3b}$), 3.05 - 3.17 (2H, m, $SCH\underline{H}'CH_{3a}$ & $SCH\underline{H}'CH_{3b}$), 3.44 (2H, s, $H\text{-}2_{Ba}$ & $H\text{-}2_{Bb}$), 3.53 - 3.84 (20H, m), 3.86 - 3.94 (2H, m, $H\text{-}6'_{Aa}$ & $H\text{-}6'_{Ab}$), 4.18 - 4.23 (1H, m, $H\text{-}1_{Ab}$), 4.32 - 4.39 (2H, m, $H\text{-}1_{Ba}$ & $H\text{-}1_{Bb}$), 4.46 - 4.51 (1H, m, $H\text{-}1_{Aa}$); δ_C (100.5 MHz, D_2O) 6.5 (q, $SCH_2\underline{C}H_{3a}$ & $SCH_2\underline{C}H_{3b}$), 39.5 (t, $S\underline{C}H_2CH_{3a}$), 40.2 (t, $S\underline{C}H_2CH_{3b}$), 60.0, 61.0, 61.0, 67.5, 68.5, 68.9, 70.9, 70.9, 72.5, 72.5, 75.3, 75.3, 75.4, 75.6, 77.0, 77.2, 79.0, 79.6 (20 x C), 87.5 (d, $C\text{-}1_{Ab}$), 90.3 (d, $C\text{-}1_{Aa}$), 102.8 (d, $C\text{-}1_{Ba}$ & $C\text{-}1_{Bb}$); HRMS (ESI-TOF): calcd. for $C_{14}H_{27}O_{11}S^+$: 403.1269. Found: 403.1284 (MH^+).

Separation of Sulfoxide Diastereoisomers by HPLC

The diastereoisomeric mixtures of *gluco* glycosyl sulfoxides **5.25** and *galacto* glycosyl sulfoxides **5.42** were separated by HPLC (column: Luna 5u C18 (100 Å) column (Phenomenex); eluent: 0.05% TFA in water; column oven: 40 °C; flow rate: 3.5 mL/min; detection: UV 210 nm) to afford single diastereoisomers:

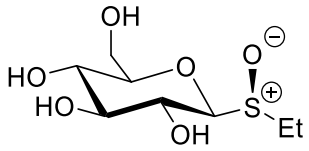
Ethyl 1-(*R*)-sulfinyl-β-D-glucopyranoside 5.25 R_S



t_R = 12.0 min; δ_H (400 MHz, D_2O) 1.19 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.82 - 2.93 (1H, m, $SC\underline{H}H'CH_3$), 3.05 - 3.17 (1H, m, $SCH\underline{H}'CH_3$), 3.37 (1H, t, *J* 9.4 Hz, H-4), 3.46 - 3.55 (2H, m, H-3 & H-5), 3.60 (1H, t, *J* 9.4 Hz, H-2), 3.69 (1H, dd, $J_{6,6'}$ 12.5 Hz, $J_{5,6}$ 4.7 Hz, H-6), 3.83 (1H,

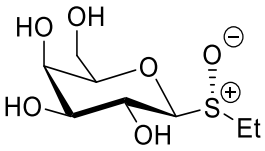
dd, $J_{6,6'}$ 12.5 Hz, $J_{5,6'}$ 1.6 Hz, H-6'), 4.16 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O) 6.5 (q, $SCH_2\underline{C}H_3$), 40.1 (t, $S\underline{C}H_2CH_3$), 60.4 (t, C-6), 67.7 (d, C-2), 68.6 (d, C-4), 76.9 (d, C-3), 80.2 (d, C-5), 87.7 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0563 (MNa^+).

Ethyl 1-(*S*)-sulfinyl-β-D-glucopyranoside 5.25 *S*S



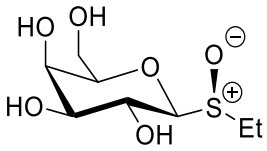
t_R = 10.9 min; δ_H (400 MHz, D_2O) 1.23 (3H, t, *J* 7.4 Hz, $SCH_2C\underline{H}_3$), 2.81 - 2.93 (1H, m, $SC\underline{H}H'CH_3$), 3.04 - 3.15 (1H, m, $SCH\underline{H}'CH_3$), 3.32 (1H, t, *J* 9.0 Hz, H-4), 3.41 - 3.53 (2H, m, H-3 & H-5), 3.55 - 3.68 (2H, m, H-2 & H-6), 3.82 (1H, d, $J_{6,6'}$ 12.5 Hz, H-6'), 4.46 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O) 6.5 (q, $SCH_2\underline{C}H_3$), 39.4 (t, $S\underline{C}H_2CH_3$), 60.7 (t, C-6), 68.8 (d, C-4), 69.1 (d, C-2), 77.1 (d, C-3), 80.7 (d, C-5), 90.4 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0563 (MNa^+).

Ethyl 1-(*R*)-sulfinyl-β-D-galactopyranoside 23 *R*S



t_R = 6.9 min; δ_H (400 MHz, D_2O) 1.20 (3H, t, *J* 7.6 Hz, $SCH_2C\underline{H}_3$), 2.85 - 2.96 (1H, m, $SC\underline{H}H'CH_3$), 3.07 - 3.18 (1H, m, $SCH\underline{H}'CH_3$), 3.62 - 3.70 (2H, m, H-5 & H-6), 3.72 – 3.80 (2H, s, H-3 & H-6'), 3.84 - 3.92 (2H, m, H-2 & H-4), 4.09 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O) 6.5 (q, $SCH_2\underline{C}H_3$), 40.2 (t, $S\underline{C}H_2CH_3$), 61.2 (t, C-6), 65.0 (d, C-2), 68.6 (d, C-3), 73.9 (d, C-4), 79.9 (d, C-5), 88.4 (d, C-1), HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0554 (MNa^+).

Ethyl 1-(*S*)-sulfinyl-β-D-galactopyranoside 23 *S*S



t_R = 7.8 min; δ_H (400 MHz, D_2O) 1.25 (3H, t, *J* 7.6 Hz, $SCH_2C\underline{H}_3$), 2.82 - 2.94 (1H, m, $SC\underline{H}H'CH_3$), 3.05 - 3.16 (1H, m, $SCH\underline{H}'CH_3$), 3.60 - 3.68 (2H, m, H-5 & H-6), 3.69 - 3.76 (2H, m, H-3 & H-6'), 3.80 (1H, t, *J* 9.6 Hz, H-2), 3.90 (1H, d, $J_{4,5}$ 3.1 Hz, H-4), 4.41 (1H, d, $J_{1,2}$ 9.8 Hz, H-1); δ_C (100.5 MHz, D_2O) 6.4 (q, $SCH_2\underline{C}H_3$), 39.4 (t, $S\underline{C}H_2CH_3$), 61.0 (t, C-6), 66.5 (d, C-2), 68.8 (d, C-3), 73.9 (d, C-4), 80.2 (d, C-5), 91.1 (d, C-1); HRMS (ESI-TOF): calcd. for $C_8H_{16}O_6SNa^+$: 263.0560. Found: 263.0554 (MNa^+).

Glycosyl Sulfoxide Inhibition Studies

Enzymes

β-Glucosidase from almonds was obtained as a lyophilised powder (3.4 mg) which was dissolved in sodium phosphate buffer solution (600 µL, 0.05 M, pH 6.8).

β-Galactosidase from *Eschericia coli* was obtained as a lyophilised powder (1.68 mg) which was dissolved in a buffer solution (600 µL, pH 7.0) containing Tris HCl (10 mM), $MgCl_2$ (10 mM) and β-mercaptoethanol (1 mM). A portion of the resulting solution (10 µL) was then diluted with the above buffer solution (40 µL) to obtain an appropriate concentration of enzyme.

α-Mannosidase from jack beans was obtained as an ammonium sulphate suspension, which was diluted with sodium acetate buffer solution (400 µL, 0.1 M, pH 4.5).

i) β-Glucosidase[55]

A stock solution of 60 mM *o*-nitrophenol β-D-glucopyranoside (0.36 g, 1.2 mmol) was prepared in sodium phosphate buffer (20 mL, 0.05 M, pH 6.8). Similar stock solutions of 10 mM ethyl 1-sulfinyl-β-D-glucopyranoside **5.25**, ethyl 1-(*R*)-sulfinyl-β-D-glucopyranoside **5.25** ***R*****S** and ethyl 1-(*S*)-sulfinyl-β-D-glucopyranoside **5.25** ***S*****S** (48 mg, 0.20 mmol) were prepared in the same buffer solution (20 mL). A variety of volumes of each solution (100 – 500 µL of substrate, 0 – 500 µL of inhibitor) were mixed and the resulting solution made up to 1 mL with the previously mentioned buffer solution. The solution was then incubated at 25 °C for 10 min. Enzyme solution (1 µL, as prepared above) was added, and the change in absorbance followed at 420 nm.

ii) α-Mannosidase[56]

A stock solution of 25 mM *p*-nitrophenol α-D-mannopyranoside (0.151 g, 0.50 mmol) was prepared in sodium acetate buffer solution (20 mL, 0.1 M, pH 4.5). A similar stock solution of 40 mM ethyl 1-(*R*)-sulfinyl-β-D-mannopyranoside **5.41** (96 mg, 0.40 mmol) was prepared in the same buffer solution (10 mL). A variety of volumes of each solution (100 – 500 µL of substrate, 0 – 500 µL of inhibitor) were mixed and the resulting solutions made up to 1 mL with the previously mentioned buffer solution. The solution was then incubated at 37 °C for 10 min. Enzyme solution (2 µL, as prepared above) was added, and the change in absorbance followed at 405 nm.

iii) β-Galactosidase[57]

A stock solution of 1.2 mM *o*-nitrophenol β-D-galactopyranoside (36 mg, 0.12 mmol) was prepared in sodium phosphate buffer solution (100 mL, 0.1 M, pH 7.0) containing 0.15 M NaCl. Similar stock solutions of 10 mM ethyl 1-sulfinyl-β-D-galactopyranoside **5.42**, ethyl

1-(*R*)-sulfinyl-β-D-galactopyranoside **5.42 *R*s** and ethyl 1-(*R*)-sulfinyl-β-D-galactopyranoside **5.42 *S*s** (48 mg, 0.20 mmol) were prepared in the same buffer solution (20 mL). A variety of volumes of each solution (100 – 500 µL of substrate, 0 – 500 µL of inhibitor) were mixed and the resulting solution made up to 1 mL with the previously mentioned buffer solution. The solution was then incubated at 37 °C for 10 min. Enzyme solution (1 µL, as prepared above) was added, and the change in absorbance followed at 420 nm.

References

1. Isobe, T.; Ishikawa, T., *J. Org. Chem.* **1999,** *64* (19), 6984-6988.

2. Mathias, E.; Shimanski, M., *J. Chem. Soc., Chem. Commun.* **1981,** (11), 569-570.

3. Gaumont, A. C.; Wazneh, L.; Denis, J. M., *Tetrahedron* **1991,** *47* (27), 4927-4940.

4. Lee, Y. C.; Stowell, C. P.; Krantz, M. J., *Biochemistry* **1976,** *15* (18), 3956-3963.

5. Pearce, O. M. T.; Fisher, K. D.; Humphries, J.; Seymour, L. W.; Smith, A.; Davis, B. G., *Angew. Chem., Int. Ed.* **2005,** *44* (7), 1057-1061.

6. Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G., *Org. Biomol. Chem.* **2003,** *1* (21), 3642-3644.

7. Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G., *Angew. Chem., Int. Ed.* **2004,** *43* (7), 828-833.

8. Chalker, J. M. M., *Angew. Chem., Int. Ed. 51* (8), 1835-1839.

9. Lin, Y. A.; Chalker, J. M.; Davis, B. G., *J. Am. Chem. Soc.* **2010,** *132* (47), 16805-16811.

10. Tanaka, T.; Huang, W. C.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Tetrahedron Lett.* **2009,** *50* (18), 2154-2157.

11. Rakotomanomana, N.; Lacombe, J.-M.; Pavia, A. A., *Carbohydr. Res.* **1990,** *197* (0), 318-323.

12. Watt, G. M.; Lund, J.; Levens, M.; Kolli, V. S. K.; Jefferis, R.; Boons, G.-J., *Chem. Biol.* **2003,** *10* (9), 807-814.

13. Ito, H.; Kamachi, T.; Yashima, E., *Chem. Commun.* **2012,** *48* (45), 5650-5652.

14. Black, W. A. P.; Colquhoun, J. A.; Dewar, E. T., *Carbohydr. Res.* **1967,** *5* (3), 362-365.

15. Murakami, H.; Minami, T.; Ozawa, F., *J. Org. Chem.* **2004,** *69* (13), 4482-4486.

16. Cochrane, S. A.; Huang, Z.; Vederas, J. C., *Org. Biomol. Chem.* **2013,** *11* (4), 630-639.

17. Yamada, S.; Sudo, A.; Goto, M.; Endo, T., *RSC Advances* **2014,** *4* (56), 29890-29896.

18. Boal, A. K.; Guryanov, I.; Moretto, A.; Crisma, M.; Lanni, E. L.; Toniolo, C.; Grubbs, R. H.; O'Leary, D. J., *J. Am. Chem. Soc.* **2007,** *129* (22), 6986-6987.

19. Maleczka Jr, R. E.; Ghosh, B.; Gallagher, W. P.; Baker, A. J.; Muchnij, J. A.; Szymanski, A. L., *Tetrahedron* **2013,** *69* (19), 4000-4008.

20. Daniel, D.; Middleton, R.; Henry, H. L.; Okamura, W. H., *J. Org. Chem.* **1996,** *61* (16), 5617-5625.

21. Trost, B. M.; Xie, J., *J. Am. Chem. Soc.* **2006,** *128* (18), 6044-6045.

22. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J., *Angew. Chem., Int. Ed.* **2014,** *53* (44), 1907-11911.

23. Schwarz, M. K.; Tumelty, D.; Gallop, M. A., *J. Org. Chem.* **1999,** *64* (7), 2219-2231.

24. Noguchi, M.; Tanaka, T.; Gyakushi, H.; Kobayashi, A.; Shoda, S. I., *J. Org. Chem.* **2009,** *74* (5), 2210-2212.

25. Tomabechi, Y.; Squire, M. A.; Fairbanks, A. J., *Org. Biomol. Chem.* **2014,** *12* (6), 942-955.

26. Li, W. S.; Zhang, N.; Sayre, L. M., *Tetrahedron* **2001,** *57* (21), 4507-4522.

27. McManus, J. M.; Herbst, R. M., *J. Org. Chem.* **1959,** *24* (8), 1042-1044.

28. Weng, S.-S.; Lin, Y.-D.; Chen, C.-T., *Org. Lett.* **2006,** *8* (24), 5633-5636.

29. France, R. R.; Compton, R. G.; Davis, B. G.; Fairbanks, A. J.; Rees, N. V.; Wadhawan, J. D., *Org. Biomol. Chem.* **2004,** *2* (15), 2195-2202.

30. France, R. R.; Rees, N. V.; Wadhawan, J. D.; Fairbanks, A. J.; Compton, R. G., *Org. Biomol. Chem.* **2004,** *2* (15), 2188-2194.

31. Kremer, C. B., *J. Am. Chem. Soc.* **1942,** *64* (6), 1285-1286.

32. Vibhute, A. M.; Muvvala, V.; Sureshan, K. M., *Angew. Chem., Int. Ed.* **2016,** *55* (27), 7782-7785.

33. Mukhopadhyay, B.; Kartha, K. P. R.; Russell, D. A.; Field, R. A., *J. Org. Chem.* **2004,** *69* (22), 7758-7760.

34. Valerio, S.; Iadonisi, A.; Adinolfi, M.; Ravidà, A., *J. Org. Chem.* **2007,** *72* (16), 6097-6106.

35. Meng, B.; Zhu, Z.; Baker, D. C., *Org. Biomol. Chem.* **2014,** *12* (28), 5182-5191.

36. Blom, P.; Ruttens, B.; Van Hoof, S.; Hubrecht, I.; Van der Eycken, J.; Sas, B.; Van hemel, J.; Vandenkerckhove, J., *J. Org. Chem.* **2005,** *70* (24), 10109-10112.

37. Li, X.; Huang, L.; Hu, X.; Huang, X., *Org. Biomol. Chem.* **2009,** *7* (1), 117-127.

38. Vic, G.; Hastings, J. J.; Howarth, O. W.; Crout, D. H. G., *Tetrahedron: Asymmetry* **1996,** *7* (3), 709-720.

39. Fried, J.; Walz, D. E., *J. Am. Chem. Soc.* **1949,** *71* (1), 140-143.

40. Purves, C. B., *J. Am. Chem. Soc.* **1929,** *51* (12), 3619-3627.

41. Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S., *Nat. Chem. Biol.* **2011,** *7* (10), 685-691.

42. Brachet, E.; Brion, J.-D.; Alami, M.; Messaoudi, S., *Chem. Eur. J.* **2013,** *19* (45), 15276-15280.

43. Tanaka, T.; Matsumoto, T.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i., *Chem. Lett.* **2009,** *38* (5), 458-459.

44. Saibal Kumar, D.; Roy, N., *Carbohydr. Res.* **1996,** *296* (1), 275-277.

45. Wang, C.; Sanders, B.; Baker, D. C., *Can. J. Chem.* **2011,** *89* (8), 959-963.

46. Saksena, R.; Zhang, J.; Kováč, P., *J. Carbohyd. Chem.* **2002,** *21* (6), 453.

47. Liu, Z.; Byun, H.-S.; Bittman, R., *Org. Lett.* **2010,** *12* (13), 2974-2977.

48. Ness, R. K.; Fletcher, H. G., Jr.; Hudson, C. S., *J. Am. Chem. Soc.* **1951,** *73*, 3742-4.

49. Mandal, S. S.; Liao, G.; Guo, Z., *RSC Advances* **2015,** *5* (30), 23311-23319.

50. Lemieux, R. U., *Can. J. Chem.* **1951,** *29* (12), 1079-1091.

51. Whitfield, D. M.; Ruzicka, C. J.; Carver, J. P.; Krepinsky, J. J., *Can. J. Chem.* **1987,** *65* (4), 693-703.

52. Wolfrom, M. L.; Cramp, W. A.; Horton, D., *J. Org. Chem.* **1965,** *30* (9), 3056-3058.

53. Hough, L.; Taha, M. I., *J. Chem. Soc.* **1956,** (0), 2042-2048.

54. Tomoo, T.; Kondo, T.; Abe, H.; Tsukamoto, S.; Isobe, M.; Goto, T., *Carbohydr. Res.* **1996,** *284* (2), 207-222.

55. Lohse, A.; Hardlei, T.; Jensen, A.; Plesner, I. W.; Bols, M., *Biochem. J.* **2000,** *349* (1), 211-215.

56. Tropea, J. E.; Molyneux, R. J.; Kaushal, G.; Pan, Y.; Mitchell, M.; Elbein, A. D., *Biochemistry* **1989,** *28* (5), 2027-2034.

57. Lo, S.; Dugdale, M. L.; Jeerh, N.; Ku, T.; Roth, N. J.; Huber, R. E., *Protein J.* **2010,** *29* (1), 26-31.